

From: Bunner, Bridget  
Sent: Tuesday, August 16, 2005 3:31 PM  
To: STIC-Biotech/ChemLib  
Subject: sequence search

Hi! I'd like to request the following sequence search for case 09/762,594 (please also include the pending databases):

1. the nucleic acid sequence of SEQ ID NO: 2
2. the nucleic acid sequence encoding the protein of SEQ ID NO: 7

Thanks!

Bridget Bunner

Art Unit 1647  
Rem 4C65  
(571) 272-0881  
mailbox 4C70

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 8/19/05  
Date Completed: 8/23/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: 1 AA#: 1 - reverse  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: OH  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 17:58:08 ; Search time 6433 Seconds  
(without alignments)  
10989.610 Million cell updates/sec

Title: US-09-762-594-2  
Perfect score: 1459  
Sequence: 1 gaattcgccgcgcgtcgac.....catgacgttaatttccttt 1459

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_cm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1339	91.8	1724	10 AF022770	AF022770 Mus muscu
2	1337.4	91.7	3405	10 BC060602	BC060602 Mus muscu
3	1203.2	82.5	3461	10 AV336075	AV336075 Rattus no
4	1200.4	82.3	1927	10 BC083877	BC083877 Rattus no
5	1079.6	74.0	2140	9 AK025520	AK025520 Homo sapi
6	1079.6	74.0	3598	9 AB043587	AB043587 Homo sapi
7	1078	73.9	3358	9 BC060792	BC060792 Homo sapi
8	1073.2	73.6	1481	6 BD261678	BD261678 Neurotran
9	1055	72.3	3572	9 AY150218	AY150218 Homo sapi
10	890	61.0	3037	9 BC034563	BC034563 Homo sapi
11	865.8	59.3	3492	5 AJ720620	AJ720620 Gallus ga
12	859	58.9	1598	4 AY644721	AY644721 Sus scrof
13	760	52.1	961	6 AX677275	AX677275 Sequence
14	691.6	47.4	859	6 CQ714948	CQ714948 Sequence
15	634.8	43.5	2714	6 BD242865	BD242865 Secretd
16	519.4	35.6	789	6 CQ731146	CQ731146 Sequence
17	287.2	19.7	18140	10 AF501319	AF501319 Mus muscu
18	287.2	19.7	193192	10 AC121292	AC121292 Mus muscu
19	280	19.2	593	10 BC050121	BC050121 Mus muscu

20	273	18.7	230768	2	AC140763	AC140763 Rattus no
21	273	18.7	240225	2	AC126290	AC126290 Rattus no
22	268.2	18.4	278	6	AX524297	AX524297 Sequence
23	268.2	18.4	278	6	AX553035	AX553035 Sequence
24	238	16.3	99251	9	AL592045	AL592045 Human DNA
25	238	16.3	140409	2	AC044825	AC044825 Homo sapi
26	238	16.3	232180	2	AC021883	AC021883 Homo sapi
27	211.4	14.5	2235	9	AK057118	AK057118 Homo sapi
28	208.2	14.3	1897	3	AK116435	AK116435 Ciona int
29	204.4	14.0	934	6	AX339070	AX339070 Sequence
30	194.2	13.3	88723	5	BX510322	BX510322 Zebrafish
31	161.8	11.1	992	5	BC054676	BC054676 Danio rer
32	157.4	10.8	1759	5	AJ720156	AJ720156 Gallus ga
33	150.8	10.3	3547	6	AX833708	AX833708 Sequence
34	150.8	10.3	3547	9	AK095650	AK095650 Homo sapi
35	150	10.3	860	6	CQ720008	CQ720008 Sequence
36	150	10.3	882	6	AX400067	AX400067 Sequence
37	150	10.3	978	9	AY598329	AY598329 Homo sapi
38	140	9.6	178951	5	BX323819	BX323819 Zebrafish
39	120.2	8.2	198133	2	AC132609	AC132609 Mus muscu
40	117.4	8.0	233018	2	AC113788	AC113788 Rattus no
41	113.4	7.8	527	6	AX339069	AX339069 Sequence
42	113	7.7	218604	2	AC111992	AC111992 Rattus no
43	110.6	7.6	554	6	CQ072121	CQ072121 Sequence
44	110.6	7.6	554	6	CQ102805	CQ102805 Sequence
45	110.6	7.6	554	6	CQ141689	CQ141689 Sequence

## ALIGNMENTS

RESULT 1	AF022770	1724 bp	mRNA	linear	ROD 13-SEP-2002
LOCUS	Mus musculus peripheral benzodiazepine receptor associated protein (Pap7) mRNA, complete cds.				
DEFINITION	AF022770				
ACCESSION	AF022770.3	GI:22831366			
VERSION					
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1724)				
AUTHORS	Li, H., Degenhardt, B., Tobin, D., Yao, Z. X., Tasken, K. and Papadopoulos, V.				
TITLE	Identification, localization, and function in steroidogenesis of PAP7: a peripheral-type benzodiazepine receptor- and PKA (Rialpha)-associated protein				
JOURNAL	Mol. Endocrinol. 15 (12), 2211-2228 (2001)				
MEDLINE	21588728				
PUBMED	11731621				
REFERENCE	2 (bases 936 to 1490)				
AUTHORS	Li, H. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-SEP-1997) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REFERENCE	3 (bases 1 to 1724)				
AUTHORS	Li, H. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-APR-2000) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REMARK	Sequence update by submitter				
REFERENCE	4 (bases 1 to 1724)				
AUTHORS	Li, H., Liu, J. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-SEP-2002) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REMARK	Sequence update by submitter				
COMMENT	On Sep 13, 2002 this sequence version replaced gi:7545290.				

## FEATURES

Source Location/Qualifiers  
1..1724  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
1..1724  
/gene="Pap7"  
61..1638  
/gene="Pap7"  
/note="PBR associated protein; PAP7"  
/codon\_start=1  
/product="peripheral benzodiazepine receptor associated protein"  
/protein\_id="AAB71197.3"  
/db\_xref="GI:22831367"  
/translation="MAAQLNVQLVSLDGLTSPDSBERPGAEGAPPTPPSSAPGN  
GLGSGAGQORPFGAAAEBAERMBQHWGFLGLELYGLALFPYKIKDQAFHPT  
YEKLKFLVHLKQVLLGPYNPDTSPEVGFDFVLGNDRREWAALGNMSKEDAMVEFK  
LLNKCPLLSAVASHRIKEBEERKRKAERREBERREERLQKEBEKRKEBEDR  
LRREBERRIRBEERLRLDEQKQIIMALNSQTAVFOQYAAQYQPNYBOQIILIRQ  
LOEHYQYMOOLYQVLAQQAALQKQOEVVMAGASLPASSKVNNTAGASDTLSYNGQ  
AKHTHENSEKVLPEAAEBALENGPKDSLPIAAPSMMTRPOIKDFKIKIRODADSVI  
TVRREENVTVRPVTHEEGSYLFWFEATDSYDIGFVYFEWTDSPNAAVSVHVSBSDE  
BEERENVTCBEKAKNKANKPLLDLPIVPYVRDCHVEVYAGSHQYPGRGVLLKFDNS  
YSLMRKSIVYRYVYTR"

## ORIGIN

Query Match 91.8%; Score 1339; DB 10; Length 1724;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1348; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 97 AGTAAAGATGGCAAGCCCTTCATCCAACTTATGAAGAAAACTGAAGTTCGTGGCACT 156  
DB 336 AATAAAGATGGCAAGCCCTTCATCCAACTTATGAAGAAAACTGAAGTTCGTGGCACT 395  
QY 157 GCATAGCAAGTCTTTTGGGCCCATATAACCCAGACAGTCCCTGAGGTGGATTCTT 216  
DB 396 GCATAGCAAGTCTTTTGGGCCCATATAACCCAGACAGTCCCTGAGGTGGATTCTT 455  
QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAAATGTCCTCAAGGA 276  
DB 456 TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAAATGTCCTCAAGGA 515  
QY 277 GATGCGCATGGTAGAGTTGTGAAGCTTCTAAATAGTGTGTCCTCTCTCTCGGCATA 336  
DB 516 GATGCGCATGGTAGAGTTGTGAAGCTTCTAAATAGTGTGTCCTCTCTCTCGGCATA 575  
QY 337 TGTTCGCTCCACAGATAG 396  
DB 576 TGTTCGCTCCACAGATAG 635  
QY 397 GCGAAGGACGCGTGAAG 456  
DB 636 GCGAAGGACGCGTGAAG 695  
QY 457 GCGAG 516  
DB 696 GCGAG 755  
QY 517 GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCCGACACTGC 576  
DB 756 GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCCGACACTGC 815  
QY 577 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCCAGGGAACACTACGAACACACAGCAGAT 636  
DB 816 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCCAGGGAACACTACGAACACACAGCAGAT 875  
QY 637 TCTCATCCGCCAGCTGCGAGGACAGCATATCAGCAGTATATAACACCGAGGAGAGAGAAC 696  
DB 876 TCTCATCCGCCAGCTGCGAGGACAGCATATCAGCAGTATATGCGAGCAGTTATATCAAGT 935  
QY 697 CCAACTGTCACAAACACAGCGCAGCATTAACAGAAACAGCAAGAGTAGTGTGCTGGGCG 756

DB 936 CCAGCTTGCAACAACACAGGCGAGCATTAACAGAAACAGCAAGAGTAGTAGTGCTGGGCG 995  
QY 757 ATCATTTGCTGTCATCATCAAAAGGTGAACACACAGCTGGAGCAAGTGTATACACTGTGAGTTAA 816  
DB 996 ATCATTTGCTGTCATCATCAAAAGGTGAACACACAGCTGGAGCAAGTGTATACACTGTGAGTTAA 1055  
QY 817 TGGACAGGCGCAAAACCCACACTGAAAAATTCGAAAAAGTCCCTTTGAGCCAGAAAGCTGCAGA 876  
DB 1056 TGGACAGGCGCAAAACCCACACTGAAAAATTCGAAAAAGTCCCTTTGAGCCAGAAAGCTGCAGA 1115  
QY 877 AGAAGCCTTTGGAAAAATGGACCAAAAGACTCTCTCCAGTGATTTGCAGCTCCATCCATGTG 936  
DB 1116 AGAAGCCTTTGGAAAAATGGACCAAAAGACTCTCTCCAGTGATTTGCAGCTCCATCCATGTG 1175  
QY 937 GACAAAGCCACAAAATCAAAGACTTTAAAGAGAAAGATTCCGAGGAGTGCAGATTCTGTGAT 996  
DB 1176 GACAAAGCCACAAAATCAAAGACTTTAAAGAGAAAGATTCCGAGGAGTGCAGATTCTGTGAT 1235  
QY 997 TACAGTAGTTCAGGAGAAAGTCTCACCGTCCGAGTCCCGACTCATGAGGAAAGGATCATATA 1056  
DB 1236 TACAGTAGTTCAGGAGAAAGTCTCACCGTCCGAGTCCCGACTCATGAGGAAAGGATCATATA 1295  
QY 1057 CCTATTTTGGGAATTTGCCACAGACAGATTATGACATTGGGTTTGGGGTTTATTTTGAATG 1116  
DB 1296 CCTATTTTGGGAATTTGCCACAGACAGATTATGACATTGGGTTTGGGGTTTATTTTGAATG 1355  
QY 1117 GACAGACTCTCCAAATGCTGTGTCAGTGTGATGTCATGTCAGTGCAGTCCAGTGCAGGAGGA 1176  
DB 1356 GACAGACTCTCCAAATGCTGTGTCAGTGTGATGTCATGTCAGTGCAGTCCAGTGCAGGAGGA 1415  
QY 1177 GGAGAGAGAGAAATGTCACTTGTGAAGAAAAACCAAAAAAGACGCCAACAGCCTCT 1236  
DB 1416 GGAGAGAGAGAAATGTCACTTGTGAAGAAAAACCAAAAAAGACGCCAACAGCCTCT 1475  
QY 1237 GCTGATGAGATTGTACCTGTACCGCGGGAGCTGTACAGAGGAAGTATATGAGGCGAG 1296  
DB 1476 GCTGATGAGATTGTACCTGTACCGCGGGAGCTGTACAGAGGAAGTATATGAGGCGAG 1535  
QY 1297 CCACAGATATCCAGGAGGGAGTCTATCTCTCAAGTTTGATAATCTCTACTCTCTGTG 1356  
DB 1536 CCACAGATATCCAGGAGGGAGTCTATCTCTCAAGTTTGATAATCTCTACTCTCTGTG 1595  
QY 1357 GAGTTCAGTCCGCTCTACTACAGAGTCTATATATCTAGATAGAGCTGCTGTTCAGAGT 1416  
DB 1596 GAGTTCAGTCCGCTCTACTACAGAGTCTATATATCTAGATAGAGCTGCTGTTCAGAGT 1655  
QY 1417 CCGGAGTCCAGGTTGAGCACAAATGACGTTTAAATTTCCCTTT 1459  
DB 1656 CCGGAGTCTAGGTTGAGCACAAATGACGTTTAAATTTCCCTTT 1698

## RESULT 2

BC060602  
LOCUS BC060602  
DEFINITION Mus musculus acyl-Coenzyme A binding domain containing 3, mRNA  
(cDNA clone MGC:79166 IMAGE:5706182), complete cds.  
ACCESSION BC060602  
VERSION BC060602.1  
KEYWORDS GI:38197725  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3405)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,



Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalil, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3405)

Strausberg, R.

Direct Submission

Submitted (31-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgaps-1@mail.nih.gov](mailto:cgaps-1@mail.nih.gov)

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>

Contact: [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu); [tom-casavant@uiowa.edu](mailto:tom-casavant@uiowa.edu)

Bonaldi, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: Plate: Row: Column: 0.

#### FEATURES

source

```
1. .3405
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:79166 IMAGE:5706182"
/tissue_type="Brain, mouse 15.5 dpc"
/clone_lib="NIH_BMAP_EX0"
/lab_host="DH10B"
/notes="Vector: pYX-ASC"
```

#### gene

```
1. .3405
/gene="Acbd3"
/notes="synonyms: 60kDa, GCP60, GOLPH1, Pap7"
/db_xref="LocusID:170760"
/db_xref="MGI:2181074"
15. .1592
/gene="Acbd3"
/codon_start=1
```

#### CDS

```
/product="acyl-Coenzyme A binding domain containing 3"
/protein_id="AAH60602.1"
/db_xref="GI:38197726"
/db_xref="LocusID:170760"
/db_xref="MGI:2181074"
/translation="MAAQLNVQLVSLDGLTLSPDSEEPGAEAGPPTPPSSAPGN
GLGSGAGGQREGEAAAGAAEARMQHGFGLLELYGLALRYKIKDGKARHPT
YEEKLFVHLKQVLGYPNPTSPFVGLGNDNRREWALGNMSKEDAMVEFVK
LNNKCCPLLSAYASERHRIEKEEERKAEERQRQEEERERLQKEEERKREBEDR
LRREERERIRIEERLQKQKQIIMAAINSOTAVOFQYAAQYPGNYEQOILIRQ
LOBHYVQYMOQLYVQLAQQAALQKQEVVMAGASLPASSKVNVTGASDLVNGQ
AKHTENSEKYLPEAEALENGPKDSLPIVAAAPSMWTRPQIKDPEKIRQDASVI
TVRGGVTVTRVFTPEEGSYLFWFATDSIDIGFVGFVFTWDSFNAASVHVSESSD
EEEEENVTCEKAKKNANKPLLDLEIVPVYVRDCHVEVAGVSHQVPGRGVYLLKFDNS
YSLWRSKSVYRYVYTR"
```

#### ORIGIN

```
Query Match          91.7%; Score 1337.4; DB 10; Length 3405;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1347; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY  97 AGTAAAGATGGCAAGCCCTTTTCATCAACTTATGAAGAAAAAAGTGAAGTTCGTGGCACT 156
Db  290 AATAAAGATGGCAAGCCCTTTTCATCAACTTATGAAGAAAAAAGTGAAGTTCGTGGCACT 349

QY  157 GCATAGCAGAGTTCTTTTGGGCCCATATACCAGACAGTCCCTCGAGTTCGATTCCT 216
Db  350 GCATAAGCAAGTTCTTTTGGGCCCATATACCAGACAGTCCCTCGAGTTCGATTCCT 409

QY  217 TGATGTGTTGGGGAATGATAGGAGAGAGAAATGGGCAGCTCTGGGAAACAATGTCCCAAGGA 276
Db  410 TGATGTGTTGGGGAATGATAGGAGAGAGAAATGGGCAGCTCTGGGAAACAATGTCCCAAGGA 469

QY  277 GGATGCGCATGTTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCGGCATA 336
Db  470 GGATGCGCATGTTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCGGCATA 529

QY  337 TGTTCGCTCCACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
Db  530 TGTTCGCTCCACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 589

QY  397 GCGAAGGCGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
Db  590 GCGAAGGCGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649

QY  457 GCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516
Db  650 GCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 709

QY  517 GAGGCTTCGCTGGAACACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576
Db  710 GAGGCTTCGCTGGAACACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 769

QY  577 CGTGCAATTCACAGCAGTATGACAGCCAGCAGTATCCAGGGAAGTACGAAACAACAGCAGAT 636
Db  770 CGTGCAATTCACAGCAGTATGACAGCCAGCAGTATCCAGGGAAGTACGAAACAACAGCAGAT 829

QY  637 TCTCATTCGCCAGCTGACAGGAGCAGCACTATCAGCAGTATATAACACAGGAGAGAGAGAGAG 696
Db  830 TCTCATTCGCCAGCTGACAGGAGCAGCACTATCAGCAGTATATAACACAGGAGAGAGAGAG 889

QY  697 CCAACTGCACAAACAACAGGAGCAGCATTACAGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAG 756
Db  890 CCAGCTTGCACAAACAACAGGAGCAGCATTACAGAAACAGCAGAGAGAGAGAGAGAGAGAGAG 949

QY  757 ATCATTTGCTGCATCATCAAGAGTGAACACAGCTGGAGCAAGTGATACACTGTCAGTTAA 816
Db  950 ATCATTTGCTGCATCATCAAGAGTGAACACAGCTGGAGCAAGTGATACACTGTCAGTTAA 1009

QY  817 TGGACAGGCGCAAAACCCACACTGAAAAATTCGAAAAAGTCTTTGAGCCAGAGAGCTGCAGA 876
Db  1010 TGGACAGGCGCAAAACCCACACTGAAAAATTCGAAAAAGTCTTTGAGCCAGAGAGCTGCAGA 1069

QY  877 AGAAGCTTTGAAAAATGACCAAAAGAGACTCTCTTCAGTGAATTCAGAGTCCATCCATGTG 936
Db  1070 AGAAGCTTTGAAAAATGACCAAAAGAGACTCTCTTCAGTGAATTCAGAGTCCATCCATGTG 1129

QY  937 GACAAGACCACAAATCAAGAGCTTTTAAAGAGAGAGATTCGGCAGGAGTGCAGATTCCTGTAT 996
Db  1130 GACAAGACCACAAATCAAGAGCTTTTAAAGAGAGAGATTCGGCAGGAGTGCAGATTCCTGTAT 1189

QY  997 TACAGTACGTCGAGGAGAGAGTCTGTCACCGTCCGAGTCCCGACTCATGAGAGAGAGATCATA 1056
Db  1190 TACAGTACGTCGAGGAGAGAGTCTGTCACCGTCCGAGTCCCGACTCATGAGAGAGAGATCATA 1249

QY  1057 CCTATTTGGGAATTTGCCACAGACAGATTATGACATTTGGGTTTGGGTTTATTTTGAATG 1116
Db  1250 CCTATTTGGGAATTTGCCACAGACAGATTATGACATTTGGGTTTGGGTTTATTTTGAATG 1309
```



```
Db 1294 CCTCTTTGGGATTTGCCACACAGATTATGACATTTGGGTTTATTTTGAATG 1353
Qy 1117 GACAGACTCTCCAAATGCTGCTGATGTCATGTCAGTGCAGTCCAGT---GACGAGGA 1173
Db 1354 GACAGACTCTCCAAACGCTGCTGATGTCATGTCAGTGCAGTGCAGTGCAGGAGGA 1413
Qy 1174 GGAGGAGGAGGAGAGAAATGTCATCTTGTGAAGAAAAGCAAAAAGAGCCCAAGCC 1233
Db 1414 GGAGGAGGAGGAGAGAAATGTCATCTTGTGAAGAGAAAGCCAAAAGATGTCAACAGCC 1473
Qy 1234 TCTGCTGGATGAGATTGTACCTGTGTACCGCGGAGCTGTACGAGAGAGATATATCCAGG 1293
Db 1474 TCTGCTGGATGAGATTGTACCTGTGTACCGCGGAGCTGTACGAGAGAGATATATCCAGG 1533
Qy 1294 CAGCCACAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGTATAATTCCTACTCTCT 1353
Db 1534 CAGCCACAGTATCCAGGAGGAGGAGGAGTCTATCTCTCAAGTTTGTATAATTCCTACTCTCT 1593
Qy 1354 GTGGAGGTCCTCAAGTCCGCTTACTACAGAGTCTATATATAGTAGAGTGTCTGTTCCAA 1413
Db 1594 GTGGAGGTCCTCAAGTCCGCTTACTACAGAGTCTATATATAGTAGAGTGTCTGTTCCAA 1651
Qy 1414 GGTCCGAGTCCAGGTTGAGCAACATGACGTTTAATTTCTTTT 1459
Db 1652 GAGTCGGAGTCTGGGTTGAGCAGAGATGACATTTAATTTCTTTT 1697
```

## RESULT 4

```
BC083877
LOCUS BC083877 1927 bp mRNA linear ROD 12-OCT-2004
DEFINITION Rattus norvegicus DMT1-associated protein, mRNA (cdna clone
MGC:95164 IMAGE:7130229), complete cds.
ACCESSION BC083877
VERSION BC083877.1 GI:54035575
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
```

## REFERENCE

```
1 (bases 1 to 1927)
Klausner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D.,
Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1927)
Director MGC Project.
Direct Submission
Submitted (01-OCT-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
```

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Howard Jacobs  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 184 Row: h Column: 2  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 33469058.

## FEATURES

```
source
1..1927
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:95164 IMAGE:7130229"
/tissue_type="Kidney, rat (Brown Norway)"
/clone_lib="NIH MGC_236"
/lab_host="DH10B"
/notes="Vector: pExpress1"
gene
1..1927
/gene="Acdb3"
/db_xref="LocusID:289312"
/db_xref="RGD:727851"
20..1600
/gene="Acdb3"
/codon_start=1
/product="DMT1-associated protein"
/protein_id="AAH83877.1"
/db_xref="GI:54035576"
/db_xref="LocusID:289312"
/db_xref="RGD:727851"
/translation="MAAOLNVQLVSLDGLTSLSPSEERPGAEGLPLOTBPSSPPGD
GLGSTAGQQRPEGAAGAAEGAEAEARMHQHWGFLGLEYLGLALRYKDKGKAFHPT
YEEKUFVALHKQVLGPNPDTSPVGFDFVLGNDRRREWAALNMSKEDAMVEFK
LNKCCPLLSAYVASHRIEKEEERKAEERREERREERREERREERREERREERRE
LRREBERRIRIEERLEQKQIMAAINSQTAVQFOYAAQYQYNGYEQOQILIRQ
LQEOHYQMOOLYVOLAAQQAQKQVAVAGASLPASAKVNTAGAGDPLPVNGQ
AKHTPENKVLPEAEAEALENGPKDSLPIVIAAPSMWTRPOIKDFEKIRODADSVI
TVRREVTVRVPTHEESYLFWEFATSDYDIFGFGYFETWDSPNAAVSVHVHSESSDD
EEEEENVTCEEKAKNVNPKLLDEIVPVYRRDCHEEYVYAGSHQYPGRGVLLKFPDN
SYSLWRSKSVYRVYTR"
ORIGIN
```

```
Query Match 82.3%; Score 1200.4; DB 10; Length 1927;
Best Local Similarity 93.4%; Pred. No. 5.2e-298; Indels 4; Gaps 2;
Matches 1276; Conservative 0; Mismatches 86;
Qy 97 AGTAAAGATGCAAGGCTTTTCATCCCACTTATGAAGAAAAAAGTGAAGTTGCGCACT 156
Db 295 AATAAAGATGCAAGGCTTTTCATCCCACTTACGAGAAAAAAGTGAAGTTGCGCACT 354
Qy 157 GCATGAAGCAAGTTCTTTTGGGCCCATATAACCCAGACAGTCCCTGAGGTTGGATTCTT 216
Db 355 GCATGAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACTTCCCTGAGGTTGGATTCTT 414
Qy 217 TCATGCTGTGGGAATGATAGGAGGAGAGAAATGGCGAGCTCTGGGAAACATGTCACAGGA 276
Db 415 TCATGCTGTGGGAATGATAGGAGGAGAGAAATGGCGAGCTCTGGGAAACATGTCACAGGA 474
Qy 277 GGATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAGTGTCTCTCTCTCTCTCTCGGCATA 336
Db 475 GGACGCCATGGTAGAGTTTGTGAAGCTTCTGAATAGTGTCTCTCTCTCTCTCAGCATA 534
Qy 337 TGTTCGTCCTCCACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
Db 535 TGTTCGTCCTCCACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594
```

```
QY 397 GCGAAGGCAGCGCTGAAGAGGAAGAACGAGAGCGGCTGCAAAAGGAGAGAGAACGGGA 456
Db 595 GCGAAGGCAGCGCTGAAGAGGAAGAACGAGAGCGGCTGCAAAAGGAGAGAGAACGGGA 654
QY 457 GCGAGAGGAGGAGGACCGGCTGAGACGGGAGGAGGAAGAGAGCGCGCGGATAGAGGAAGA 516
Db 655 GCGAGGAGGAGGAGGACCGGCTGAGACGGGAGGAGGAAGAGAGCGCGCGGATAGAGGAAGA 714
QY 517 GAGGCTTGGCTGGAAACAGCAAAAGCAGCAGATAATGCGAGCTTTAAACTCGCAGACTGC 576
Db 715 GAGGCTTGGCTGGAAACAGCAAAAGCAGCAGATAATGCGAGCTTTAAACTCGCAGACTGC 774
QY 577 CGTGAATTCAGCAGTAGTGACGCCAGCAGTAGTATCCAGGAACTACGACACACACAGAT 636
Db 775 CGTGAATTCAGCAGTAGTGACGCCAGCAGTAGTATCCAGGAACTACGACACACACAGAT 834
QY 637 TCTCATCCGCCAGCTGCGAGGAGCAGCACTATCAGCAGTATAAACCAGGAGCAGCAAC 696
Db 835 TCTCATCCGCCAGCTGCGAGGAGCAGCACTATCAGCAGTATAAACCAGTATACCAAGT 894
QY 697 CCAACTGTCACAAACAGGAGCAGATTACAGAAAACAGCAAGTAGTGATGGCTGGGCG 756
Db 895 CCAGCTTGACAGCAACAGCAGCATTGCGAGGAAGTAGTGATGGCTGGGCG 954
QY 757 ATCATTGCTGTCATCATCAAGGTTGAACACAGCTGGAGCACTGATACACTGTCAGTTAA 816
Db 955 TTCATTGCTGCTGCTGAGCAAAAGTGAACACAGCTGGAGCAGTAGTCCCTGCCAGTGA 1014
QY 817 TGGCAGGCGCAAAACCCACACTGAAAATTCGAAAAAGTCCCTTGAGCAGAGCTGCAGA 876
Db 1015 CGGACAGGCCAAGCCCACTGAGNATCCCGAGAAAGTCTTGAGCCAGAGCTGCGGA 1074
QY 877 AGAAGCTTGGAAAATGGAACCAAAAGACTCTCTTCAGTGATGTCAGCTCCATCCATGTG 936
Db 1075 AGAGGCTTGGAAAATGGAACCAAAAGACTCTCTTCAGTGATGTCAGCTCCCTCCATGTG 1134
QY 937 GACAGACCAAAATCAAGACTTTAAAGAGAGATTCGGCAGGATGAGATTCGTGAT 996
Db 1135 GACAAGACCAAGATCAAAAGACTTTAAAGAGAGATTCGGCAGGATGAGATTCGTGAT 1194
QY 997 TACAGTACGTCGAGGAGAGTTCGTACCGTCCGAGTCCCGACTATGAGGAGAGATCATA 1056
Db 1195 TACAGTCCGCGAGGAGAGTACGTACCGTTCAGTCCCGACTATGAGGAGAGATCATA 1254
QY 1057 CTAATTTTGGGAATTTGCCACAGACAGTTATGACATTTGGTTTGGGTTTATTTGAATG 1116
Db 1255 CCTCTTTTGGGAATTTGCCACAGACAGTTATGACATTTGGTTTGGGTTTATTTGAATG 1314
QY 1117 GACAGACTCTCAAAATGCTGTCAGTGTCATGTCAGTCCAGTCCAGT --GACGAGA 1173
Db 1315 GACAGACTCTCAAAACGCTGCTGTCAGTGTCATGTCAGTCCAGTCCAGTCCAGGAGAGA 1374
QY 1174 GAGAGGAGGAGGAAGAAAATGTCACTTGTGAAGAAAAGCAAAAGAGAGCCCAAGCC 1233
Db 1375 GAGAGGAGGAGGAAGAAAATGTCACTTGTGAAGAGAAAGCCAAAAGAGATGTCAACAGCC 1434
QY 1234 TCTGCTGGATGAGATTGTACCTGTGTACCGCGGAGCTGTACAGGAGAGATATATGCAAG 1293
Db 1435 TCTGCTGGATGAGATTGTACCTGTGTACCGCGGAGCTGTATGAGGAAGTGTATGCAAG 1494
QY 1294 GAGCCACAGATATCCAGGAGGGAGTCTATCTCTCAAGTTTGAATTTCTACTCTCT 1353
Db 1495 GAGCCACAGATATCCAGGAGGGAGTCTATCTCTCAAGTTTGAATTTCTACTCTCT 1554
QY 1354 GTGGAGGTCCAAAGTCCGTACTACAGAGTCTATTATACAGAGCTGTCTGTCTCCAA 1413
Db 1555 GTGGAGGTCTAAGTCAAGTCTACTACAGAGTCTATTATACAGATTAAGCTGTGTACAG 1614
QY 1414 GGTCCGAGTCCAGGGTTGAGCACAACATGACGTTTAAATTTCTCTT 1459
Db 1615 AGT-CGGAGTCTGGGGTTGAGCAGAAGATGACATTTAAATTTCTCTT 1659
```

```
RESULT 5
AK025520
LOCUS Homo sapiens cDNA: FLJ21867 fis, clone HEP02419. 2140 bp mRNA linear PRI 13-SEP-2003
DEFINITION AK025520
ACCESSION AK025520
VERSION AK025520.1 GI:10438060
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2140)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan [E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416]
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source location/Qualifiers
1..2140
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP02419"
/cell_line="HepG2"
/clone_lib="HEP"
/note="Cloning vector pME18SFL3"
49..1635
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:10438061"
/translation="MAAVNAERLEVSVDGLTSPDPERPCAEGAPLLPPPLPPSP
PGSRGPASGEQPEPGEAAAGAAAEARRLEQRWGFGLLELYGLALRFFKDKKAF
HPTVEKKLVKLVHKKQVLMGPNPDTCPEVGFDFVLGNDRRRWAALGNMSKEDAWVE
EERLRREERRRIIEERLRLEQKKQIIMAAINSQTAVQFOYAAQOYPGNYEQOQIL
TROJQEOHYQOYMOOLYOVLAQQAALQKQEVVAGSLSPTSKVNATVPSNMWSV
NQAKTHTDSSEKLEPEAAEALENGPKESLPVIAAASMTTRPOIKDFKPKIQQAD
SVITVGRGVVTVRVTHERGSLYFWEFATNDIGFGVTFWTDSPMTAVSVHVES
SDDDEEBEENIGCEBKAKNKANPLLDIEIVPVYRRDCHBEVYAGSHQYPGPGVYLLKF
DNSYSLRWSKSVYRVVYTR"
CDS
Query Match 74.0%; Score 1079.6; DB 9; Length 2140;
Best Local Similarity 87.2%; Pred. No. 7e-267;
Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
QY 97 AGTAAAGATGCGAAAGCCTTATCCCAACTTATGAAGAAAACCTGAAAGTTCTGGCACT 156
Db 333 AGAAAAGATGGCAAGCATTTATCCCACTTATGAAGAAAATTTGAAGTTGTGCACT 392
QY 157 GCATTAAGCAAGTTCTTTTGGGCCCATATAAACCCAGACAGCTCCCTAGGTTGGATTCTT 216
Db 393 GCATTAAGCAAGTTCTTTATGGGCCCATATAAATCCAGACACTTGTCTCTGAGTTGGATTCTT 452
```

QY 217 TGAATGTTGGGAATGATAGGAGGAGAGAAATCGGCGAGCTCTGGGAAAACATGTCGAAGGA 276  
 Db 453 TGAATGTTGGGAATGATAGGAGGAGAGAAATCGGCGAGCTCTGGGAAAACATGTCGAAGGA 512  
 QY 277 GGAATGCATCGTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCCTCTCTCGGCATA 336  
 Db 513 GGAATGCATCGTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCCTCTCTCGGCATA 572  
 QY 337 TGTGTGGTCCCAAGATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396  
 Db 573 TGTGTGGTCCCAAGATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 632  
 QY 397 GCGAAGCGCCTGTAAG 456  
 Db 633 GCGAAGCGCCTGTAAG 692  
 QY 457 GCGAAGAGGAG 516  
 Db 693 GAG 752  
 QY 517 GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATTAATGGCAGCTTTAAACTTCGCAGACTGC 576  
 Db 753 AAGGCTTCGGCTGGAGCAGCAAAAGCAGCAGATTAATGGCAGCTTTAAACTTCGCAGACTGC 812  
 QY 577 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCCAGGAACTACGAAACACAGCAGAT 636  
 Db 813 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCCAGGAACTACGAAACACAGCAGCAAT 872  
 QY 637 TCTCATCCCGAGCTCCAGGAGCAGCACTATCAGCAGTATTAACACACAGGAGCAAAAC 896  
 Db 873 TCTCATCCCGAGCTCCAGGAGCAGCACTATCAGCAGTATTAACACACAGGAGTGTATCAAGT 932  
 QY 697 CCAACTGTCACAAACAGCAGCAGCATTACAGAAACAGCAAGAGTATGATGCTGGGGC 756  
 Db 933 CCAGCTTGACAGCAGCAGCAGCATTACAGAAACAGCAAGAGTATGATGCTGGGGC 992  
 QY 757 ATCATTTGCTGATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCAGTTAA 816  
 Db 993 TTTCTTTGCTTACATCATCAAAAGTGAATGCACTTACCAAGTAAATATGATGTCAGTTAA 1052  
 QY 817 TGAACAGGCGCAAAACCAACACACTGAAATTCGGAAGTCTTTCAGCCAGAGCTGCAGA 876  
 Db 1053 TGAACAGGCGCAAAACCAACACACTGAAATTCGGAAGTCTTTCAGCCAGAGCTGCAGA 1112  
 QY 877 AGAAGCTTTGGAAATGGACCAAAAGACTCTTTCAGTGAATGTCAGCTCCATCCATGTC 936  
 Db 1113 AGAAGCTTTGGAAATGGACCAAAAGACTCTTTCAGTGAATGTCAGCTCCATCCATGTC 1172  
 QY 937 GACAAGACCAAAATCAAGACTTTAAAGAGAGATTCGCGAGGATTCAGATTTCTGTGAT 996  
 Db 1173 GACAAGCTTCAGATCAAAAGACTTTCAAGAGAGATTCAGCAGGATTCAGATTTCTGTGAT 1232  
 QY 997 TACAGTACGTGAGGAGAGATGTCACCGTCCAGTCCCGACTCATGAGAGAGATCATTA 1056  
 Db 1233 TACAGTGGGCGGAGGAGAGTGTGTCATCTGTCAGTACCCACCCATGAAAGAGATCATTA 1292  
 QY 1057 CCTATTTCGGAATTTGCACAGACAGTATGACATTTGGTGGTGGTGGTGGTGGTGGTGG 1116  
 Db 1293 TCTCTTTGGGAATTTGCCACAGACAAATATGATGATGCGTGGTGGTGGTGGTGGTGGTGG 1352  
 QY 1117 GACAGACTCTCCAAATGCTGCTGTCAGTGTGTCATGTCAGTGTGTCAGTGTGTCAGTGTG 1176  
 Db 1353 GACAGACTCTCCAAATGCTGTCAGTGTGTCATGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 1412  
 QY 1177 GAGGAGGAGAGAAATGTCATTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236  
 Db 1413 GAGGAGGAG 1472  
 QY 1237 GCTGGATGAGTTGTTACCTGTTGACCGGGGAGCTGTCAGGAGAGATGATGAGGAGAG 1296  
 Db 1473 GCTGGATGAGTTGTTGCTGTTGACCGGGGAGCTGTCAGGAGAGAGATGATGAGGAGAG 1532  
 QY 1297 CCACAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATATTTCTACTCTCTGTCG 1356

Db 1533 CCATCAATATCCAGGAGAGAGTCTATCTCTCAAGTTTGACAACCTCTACTCTTTGTG 1592  
 QY 1357 GAGGTCCTCAAGTCGTCCTACTACAGAGTCTATTATATCTAGATAGAGTCTGTTCCAGGT 1416  
 Db 1593 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATATCTAGATATAAAATGTTGTACAAGT 1652  
 QY 1417 CCGAGTCCAGGTTTGAGCACAACATGACGTTTAAATTT 1454  
 Db 1653 CTGAGTCTAGGTTTGGCGAGAGATGACATTAAATTT 1690  
 RESULT 6  
 AB043587 3598 bp mRNA linear PRI 28-NOV-2001  
 LOCUS Homo sapiens gcp60 mRNA for golgi resident protein GCP60, complete cds  
 DEFINITION  
 ACCESSION AB043587  
 VERSION AB043587.2 GI:15799258  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Sohda,M., Misumi,Y., Yamamoto,A., Yano,A., Nakamura,N. and Ikehara,Y.  
 TITLE Identification and characterization of a novel Golgi protein, GCP60, that interacts with the integral membrane protein giantin  
 JOURNAL J. Biol. Chem. 276 (48), 45298-45306 (2001)  
 MEDLINE 21576278  
 PUBMED 11590181  
 REFERENCE 2 (bases 1 to 3598)  
 AUTHORS Misumi Y., Ikehara, Y. and Sohda, M.  
 DIRECT SUBMISSION  
 TITLE Submitted (23-MAY-2000) Yoshio Misumi, Fukuoka University School of Medicine, Department of Biochemistry; 45-1, 7 chome Nanakuma, Jona-ku, Fukuoka, Fukuoka 814-0180, Japan  
 JOURNAL (E-mail: misumi@fukuoka-u.ac.jp, Tel:81-92-801-1011(ex.3251), Fax:81-92-864-3865)  
 COMMENT On Sep 27, 2001 this sequence version replaced gi:12060408.  
 FEATURES  
 Location/Qualifiers  
 1..3598  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="Hela"  
 /note="vector:HB814"  
 1..3598  
 /gene="gcp60"  
 56..1642  
 /gene="gcp60"  
 /codon\_start=1  
 /product="golgi resident protein GCP60"  
 /protein\_id="BAB20592.2"  
 /db\_xref="GI:15799259"  
 /translation="MAAVLNARLEVSVDGLTSLDPERPCAGAPLPPPLPPSP  
 PGRSGPGASGEQPEFGAAGAAEABRRLEQRNGFGLEELYGLALRFKDKGAF  
 HPTVEBKLLVALHKQVLMGPYNPTDCEVGFDFVLGNDRREREAALGNMSKEDAMV  
 FVLNRCVHLRSTTVASHKIEKEQEKREERERERERERERERERERERERERER  
 BERLAREER  
 IROLQEQVQYMOQLYVOLAQQQAALQKQEVVAVAGSSLPSSKVNATVPSNMVS  
 NQOAKTHYDSSEKLEPEAEALENGPKESLPVIAASPMWTRPQIKDKPEKIQODAD  
 SVITVGRGEVTVRPTHEEGSYLFWFATNDYDGFYFWFTWSPNTAVSHVSES  
 SDDDBSEENIGCEEKAKKNANKPLLDLDEIVPVYRRDCHCEVYAGSHQVPGRGVYLKLF  
 DNSYSLSKSSVYRVYTR"  
 1643..3580  
 /gene="gcp60"  
 3' UTR  
 ORIGIN  
 Query Match 74.0%; Score 1079.6; DB 9; Length 3598;  
 Best Local Similarity 87.2%; Pred. No. 7.5e-267;  
 Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 97 AGTAAAAGATGCGAAGCCCTTTTCATCCRACTTATGAAGAAACCTCAAGTTCGTGGCACT 156  
Db 340 AGAAAAAGATGCGAAGCAATTTTCATCCAACTTATGAAGAAAATTTGAAGCTTTGTGGCACT 399  
QY 157 GCATAAGCAAGTCTTTTGGGCCCATATATACCCAGACACGTCCTCCCTGAGGTGGATTCTT 216  
Db 400 GCATAAGCAAGTCTTTATGGGCCCATATATATCCAGACATTTGTCTGTGAGTTGGATTCTT 459  
QY 217 TGATGTGTTGGGAATGATAGAGAGAGAAATGGGAGCTCTGGGAAACATGTCGAAGA 276  
Db 460 TGATGTGTTGGGGAATGACAGAGAGAGAAATGGGAGCTCTGGGAAACATGTCGAAGA 519  
QY 277 GCATGTCATGTTAGAGTGTGCAAGCTTCTAAATAGTGTGCTCTCTCTCTCTCGGCATA 336  
Db 520 GCATGTCATGTTGGAGTGTGCAAGCTTCTAAATAGTGTGCTCTCTCTCTCTCTCAACATA 579  
QY 337 TGTTCGCTCCCAAGATAGAAAGGAAGAAAGAGAAAGAAAGAAAGCGGAGAGGA 396  
Db 580 TGTTCGCTCCCAAGATAGAAAGGAAGCAAGCAAGAAAGAAAGAAAGGAAGAGGA 639  
QY 397 GCGAAGGACGCTGAAGAGGAAGAACAGAGCGGCTGCAAAAGGAAGGAAGAGCGGAA 456  
Db 640 GCGAAGGCGGCTGAAGAGGAAGAAAGAAAGAAAGCTCTGCAAAAGGAGGAAGAAAGCTAG 699  
QY 457 GCGAGAGGAGGAGCCGCTGAGACGGGAGGAGAGAGAGCGCGGATAGAGGAGA 516  
Db 700 GAGAGAAAGAGGAAGAGCTTCGACGGGAGGAGAGGAAAGAGACCGGATAGAGAGA 759  
QY 517 GAGGCTTCGCTGGAACACAAAGCAGCAGCAATATGGCAGCTTTAAACTCCGACACTGC 576  
Db 760 AAGGCTTCGCTGAGCAGCAAAAGCAGCAGATATATGGCAGCTTTAAACTCCGACACTGC 819  
QY 577 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCAGGGAACACTAGCAACACAGCAGAT 636  
Db 820 CGTGCAATTCAGCAGTATGACGCCAGCAGTATCAGGGAACACTAGCAACACAGCAGCAAT 879  
QY 637 TCTCATCCGCGCTGACGAGCAGCAGTATCAGCAGTATTAACACCAGGAGCAACAAAC 696  
Db 880 TCTCATCCGCGCTGACGAGGACAACTATCAGCAGTATATGTCAGCAGCTGTATCAAGT 939  
QY 697 CCAACTGTCACAAACACGAGCAGCAATACAGAAAACAGCAAGAGTAGTATGGCTGGGGC 756  
Db 940 CCAAGCTTGACAGCAACAGCGCAGCATTACAGAAAACAGCAGGAAGTAGTAGTGGCTGGTC 999  
QY 757 ATCATGCTGTCATCATCAAGGTTGAACACAGCTGAGCAAGTATACACTCTGCTGTTAA 816  
Db 1000 TTCTTGTGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAGTTAA 1059  
QY 817 TGGACAGGCCAAACCCACACTGAAATTCGAAAAGTCCCTTGAGCCAGAGCTGCAGA 876  
Db 1060 TGGACAGGCCAAACACACTGACAGCTCCGAAAAGAACTGGAAACAGAGAGCTGCAGA 1119  
QY 877 AGAAGCTTTGAAAATGGACCAAAAGACTCTCTTCAGTGTATGAGCTCCCATCCATGTG 936  
Db 1120 AGAAGCTTTGGAATGGACCAAAAGATCTCTTCAGTATAGCAGCTCCATCCATGTG 1179  
QY 937 GACAGACCAAAATCAAGAATTTAAGAGAAGATTCGGCAGGATGCGAGTCTGTGAT 996  
Db 1180 GACAGACCTCAGATCAAAAGCTTCAAGAGAAGATTCAGCAGGATGCGAGTTCGGTGT 1239  
QY 997 TACAGTACGTCGAGAGAGTGTCCAGCTCCGAGTCCGACTCATGAGGAGGATCATA 1056  
Db 1240 TACAGTGGGCCGAGGAGAGTGTGTCTGTTCCAGTACCCACCCATGAAGAGGATCATA 1299  
QY 1057 CTTATTTTGGGAATTTGCCACAGACAGTATGACATTTGGGTTTGGGTTTATTTTGAATG 1116  
Db 1300 TCTCTTTTGGGAATTTGCCACAGACAAATATGACATTTGGGTTTGGGTTTATTTTGAATG 1359  
QY 1117 GACAGACTTCCAAATGCTGCTGTGAGTGTGATGTGATGTGAGTCCAGTACGAGAGGA 1176  
Db 1360 GACAGACTTCCAAACACTGCTGTGAGCGTGTGATGTGATGTGAGTCCAGCGATGACGAGCA 1419

QY 1177 GGAGGAGGAAGAAATGTCACTTGTGAAGAAAAAGCAAAAGAACGCCAACGCTCT 1236  
Db 1420 GGAGGAAGAAAGAAACATCGGTTGTGAAGAGAAAGCAAAAGAAATGCCAACGCTCTT 1479  
QY 1237 GCTGATGAGATTTGTACCTGTGTACCGCGGGAGCTGTCCAGGAGGAATATATGACAGGAG 1296  
Db 1480 GCTGATGAGATTTGTGCTGTGTACCGAGCGGAGCTGTCTATGAGGAGGTGTATGCTGGCAG 1539  
QY 1297 CCACAGATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATATTAATCTCTCTCTGTG 1356  
Db 1540 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGCAAACTCTCTCTCTGTG 1599  
QY 1357 GAGGTCCAAAGTCCGCTACTACAGAGTCTATTATATAGATAGAGCTGTGTTCCAAGT 1416  
Db 1600 GCGTCAAAATCAGTCTACTACAGAGTCTATTATCTAGATATAAATGTTGTACAAAGT 1659  
QY 1417 CGGAGTCCAGGTTGAGCACAAACATGACGTTTAAATTT 1454  
Db 1660 CTGGAGTCTAGGTTGGCAGAGATGACATTTAAATTT 1697  
RESULT 7  
BC060792  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
JOURNAL  
REMARK  
COMMENT  
3358 bp mRNA linear PRI 11-DEC-2003  
Homo sapiens acyl-Coenzyme A binding domain containing 3, mRNA  
(cDNA clone MGC:71593 IMAGE:4813993), complete cds.  
BC060792  
BC060792.1 GI:38511511  
MGC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.S., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Schneitz,T.E., Brownstein,M.J., Uadin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettnerman,M., Madan,A., Rodrigues,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Schnurich,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
1247932  
2 (bases 1 to 3358)  
Strausberg,R.  
Direct Submission  
Submitted (03-NOV-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>



Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 133 Row: 1 Column: 24  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15826851.

## FEATURES

## source

1..3358  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:71593 IMAGE:4813993"  
 /tissue\_type="Brain, hippocampus"  
 /clone\_lib="NIH MGC\_95"  
 /lab\_host="DH10B"  
 /notes="Vector: pBluescript"  
 1..3358  
 /gene="ACBD3"  
 /notes="synonyms: GCP60, PAP7, GOCAP1, GOLPH1"  
 /db\_xref="LocusID:64746"  
 /db\_xref="MIM:606809"  
 44..1630  
 /gene="ACBD3"  
 /codon\_start=1  
 /product="ACBD3 protein"  
 /protein\_id="AAH60792.1"  
 /db\_xref="GI:38511512"  
 /db\_xref="LocusID:64746"  
 /db\_xref="MIM:606809"  
 /translation="MAAFLNARRLEVSVDGLTSLDPPERPGAEGLPLPPPLPPSP  
 PGSGRPGAGSQPEGEAAGAAEARELRQRWGFGLLEELGLALRPFKEDKAP  
 HPTYEKLKLVALHKGMPINDTCEVGFDFVIGNDNRREARALGNMKEDAWPE  
 FVLNLRCCFLFSTVYASHKIEKKEQKKRKEERREERREERLEQKEERREE  
 EERLRREERRIEERLRLEQKQKQIIMALNSOTAVQFOYAAQOYQPNVEQOQIL  
 IRLOEQHYQQYMQOYLQVLAQQAALQKQVAVAGSLPTSSKNATVPNSMWSV  
 NQOAKHTDSSKMELEPEAAEALENGPKESLPVIAAPSMWTRPOIKDPKEKIQDAD  
 SVITVGRGVTVRVPVTHEEGSYLFWEPATDNDYDIFGVYFVFWTSPNTAVSVHSES  
 SDDDEENIGCEKAKKANRPLDLEIVPVIRDCHEEVYAGSHQYVGRGVYLLKF  
 DSNYSLSRKSYYVYYVYTR"  
 353..550  
 /gene="ACBD3"  
 /note="ACBP; Region: Acyl CoA binding protein"  
 /db\_xref="CDD:pfam00887"

## CDS

## misc\_feature

## ORIGIN

Query Match 73.9%; Score 1078; DB 9; Length 3358;  
 Best Local Similarity 87.1%; Pred. No. 1.9e-266;  
 Matches 1183; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 97 AGTAAAGATGGCAAGCCTTTCATCCAACTTATGAAGAAAACTGAAGTTCTGTCGCACT 156  
 DB 328 AGAAAAAGATGGCAAGCATTTCATCCAACTTATGAAGAAAAATGAAGCTTGTGCACT 387  
 QY 157 GCATAGCAAGTTCTTTTGGGCCCATATACCAGACAGTCCCTGAGGTTGGATTCTT 216  
 DB 388 GCATAGCAAGTTCTTTTGGGCCCATATATATCCAGACACTTGTCTGAGTTGGATTCTT 447  
 QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAATGGGAGCTCTGGGAAACATGTCGAAGGA 276  
 DB 448 TGATGTGTTGGGAATGACAGAGGAGAGAGATGGGAGCCCTGGGAAACATGTCGAAGA 507  
 QY 277 GGATGCCATGGTAGAGTTTGAAGCTTCTAAATAGTGTGTTCCTCTCCTCTCGGCATA 336  
 DB 508 GGATGCCATGGTAGAGTTTGAAGCTTCTAAATAGTGTGTTCCTCTCCTCTCCTCAACATA 567  
 QY 337 TGTTCGCTCCACAGATAG 396  
 DB 568 TGTTCGCTCCCAAAATAG 627  
 QY 397 GCGAGGCGCATGGAAG 456

628 GCGAAGCGCGCTGAAG 687  
 QY 457 GCGAG 516  
 DB 688 GAG 747  
 QY 517 GAGGCTTCGGCTGGAAACAGACAGCAATAATGGCAGCTTTAAACTTCGACAGACTGC 576  
 DB 748 AAGGCTTCGGTGGAGCAGCAAAAGCAGCAGATATGGCAGCTTTAAACTTCGACAGACTGC 807  
 QY 577 CGTGAATTCAGCAGATATGCGAGCCAGCAGTATCCAGGGAATACGAAACAAACAGCAGAT 636  
 DB 808 CGTGCAGTTTCCAGCAGTATGCGAGCCCAACAGTATCCAGGGAATACGAAACAGCAGCAAT 867  
 QY 637 TCTCATCCGCGAGCTGCGAGCAGCAGTATCAGCAGTATTAACACACAGCAGCAGCAAC 696  
 DB 868 TCTCATCCGCGAGTTGCGAGGCAACACTATCAGCAGTATCATGCGAGCTTGTATCAAGT 927  
 QY 697 CCAACCTGCACAAACACAGCGCAGCATTACAGAAAACAGCAAGAGTAGTGTGCTGGGC 756  
 DB 928 CCAGCTTGCACAGCAACAGCGCAGCATTACAGAAAACAGAGGAGTAGTAGTGGCTGGTC 987  
 QY 757 ATCATTCCTGTCATCATCAAGGTGAACACAGCTGGAGCAAGTGATACACTGTCAAGTTAA 816  
 DB 988 TTTCTTGCCTCATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAAGTTAA 1047  
 QY 817 TGGCAGGCGCAAAACCCACACTGAAAATTCGAAAAAGTCTTTCAGCAGAGAGAGCTCAGA 876  
 DB 1048 TGGCAGGCGCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCAAGAGCTCAGA 1107  
 QY 877 AGAAGCTTGGAAAAATGGACAAAAGACTCTCTCCAGTGATTGCAAGCTCCATCCATGTG 936  
 DB 1108 AGAAGCTTGGAAATGGACAAAAGAAATCTCTTCAGTAATAGCAGCTCCATCCATGTG 1167  
 QY 937 GACAAAGCACAATAACAAAGCTTAAAGAGAGAGATTCGCGAGGATGCGAGATTTCTGTGAT 996  
 DB 1168 GACAGCCTCAGATCAAAAGCTTCAAGAGAGAGATTCAGCAGGATGCGAGATTCGCTGAT 1227  
 QY 997 TACAGTACGTCCAGAGAGAGTCTGTCACCGTCCGAGTCCGAGTCCATGAGAGAGATCATA 1056  
 DB 1228 TACAGTGGGCGGAGGAGAGTGTGTCACGTTCAGTACCCACCCATGAAGAGAGATCATA 1287  
 QY 1057 CCTATTTTGGGAATTTGCCACAGACAGTTATACATGTTGGGTTTGGGTTTATTTTCAATG 1116  
 DB 1288 TCTCTTTGGGAATTTGCCACAGACAAATATGACATTTGGGTTTGGGTTTATTTTGAATG 1347  
 QY 1117 GACAGACTCTCCAAATGCTGTGTCAGTGTGTCATGTTCAGTGTGTCAGTGTGTCAGGAGGA 1176  
 DB 1348 GACAGACTCTCCAAACACTGTCTGTGTCAGCGTGTGTCAGTGTGTCAGCGATGACGACGA 1407  
 QY 1177 GGAG 1236  
 DB 1408 GGAG 1467  
 QY 1237 GCTGATGAGATTGTACCTGTGTACCGGGGAGCTGTCCAGAGGAGATATATGCGGAG 1296  
 DB 1468 GCTGATGAGATTGTGCTGTGTACCGGAGCTGTCTATGAGGAGGTTGTGCTGGCAG 1527  
 QY 1297 CCACCACTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTCTG 1356  
 DB 1528 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTCTG 1587  
 QY 1357 GAGGTCCCAAGTCCGCTCTACTACAGAGTCTATATATAGTATAGAGTGTGTTCAGAGT 1416  
 DB 1588 GCGGTCAAAATCAGTCTACTACAGAGTCTATATATAGTATAGTATAGTATAGTATAGT 1647  
 QY 1417 CCGGAGTCCAGGTTGAGCAGCAACATGAGCTTTAATTT 1454  
 DB 1648 CTTGAGTCTAGGTTGGGAGAGAGATGATTAATTT 1685

RESULT 8

BD261678	1481 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	Neurotransmission associated proteins.			
DEFINITION	BD261678			
ACCESSION	BD261678.1 GI:33071446			
VERSION	JP 2002519064-A/3.			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1481)			
AUTHORS	Lal,P., Tang,T.Y., Yue,H., Corley,N.C., Guegler,K.J., Gorgone,G.A., Baughn,M.R. and Patterson,C.			
TITLE	Neurotransmission associated proteins			
JOURNAL	Patent: JP 2002519064-A 3 02-JUL-2002;			
COMMENT	INCYTE PHARMACEUTICALS INC OS Homo sapiens (human) PN JP 2002519064-A/3 PD 02-JUL-2002 PF 02-JUL-1999 JP 2000558211 PR 02-JUL-1998 US 60/091677 PI PREETI LAL,TOM Y TANG,HENRY YUE,NEIL C CORLEY,KARL J GUEGLER, GINA A GORGONE,MARIAH R BAUGHN,CHANDRA PATTERSON PC C12N15/09,A61K38/00,A61P25/00,A61P35/00,A61P37/02,C07K14/47, PC C07K14/705 PC C07K16/18,C07K16/28,C12N5/10,C12Q1/68,C12N15/00,A61K37/02, PC C12N5/00 CC Incyte Clone No: 998868 FH Key FT source 1..1481 FT Location/Qualifiers /organism='Homo sapiens (human)'. 1..1481 Location/Qualifiers /organism='Homo sapiens (human)'. /mol_type='genomic DNA' /db_xref='taxon:9606'			
FEATURES	source			
ORIGIN	Query Match 73.6%; Score 1073.2; DB 6; Length 1481; Best Local Similarity 86.9%; Pred. No. 3e-265; Matches 1180; Conservative 0; Mismatches 178; Indels 0; Gaps 0;			
QY	97	AGTAAAGATGCAAGCCCTTTCATCCACTTATGAAGAAACCTGAAGTTGCGGCAC	156	
DB	70	AGAAAGATGCAAGCAAGCATTTATCCCACTTATGAAGAAATTTGAAGCTTGTGGC	129	
QY	157	GCATTAAGCAAGTCTTTTGGGCCCATATAACCCAGACACGTCCTCGAGGTTGGATTCT	216	
DB	130	GCATTAAGCAAGTCTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTTGGATTCT	189	
QY	217	TGATGTGTTGGGAATGATAGAGAGAGAAATGGGAGCTCTGGGAAACATGTCCAAGGA	276	
DB	190	TGATGTGTTGGGAATGACAGGAGAGAGAAATGGGAGCCCTGGGAAACATGTCTAAGA	249	
QY	277	GCATGCCATGTAGAGTTGTGAAGCTCTTAATAGTGTCTCTCTCTCTCTCTCTCTCT	336	
DB	250	GGATGCCATGTGGAGTTGTCAAGCTCTTAATAGTGTCTCTCTCTCTCTCTCTCTCT	309	
QY	337	TGTTGCGTCCCAAGATAGAGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	396	
DB	310	TGTTGCGTCCCAAGATAGAGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	369	
QY	397	CGGAAGGACGCTGAAGAGGAAGAACGAGAGCGGTGCAAGAGGAAGGAAGGAAGCGGAA	456	
DB	370	CGGAAGGCGCGTGAAGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	429	
QY	457	CGGAGAGGAGGAGCCGCTCAGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA	516	
DB	430	GAGAGAAGGAAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA	489	
QY	517	GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCGACACTGC	576	
DB	490	AAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCGACACTGC	549	

QY	577	CGTCAATTTCCAGCAGTATGCAGCCAGCAGTATCCAGGGAACTACGAACACAGCAGAT	636	
DB	550	CGTGAGTTCCAGCAGTATGCAGCCCAACAGATATCCAGGGAACTACGAACAGCAGCAAT	609	
QY	637	TCTCATCCGCGAGTGCAGGAGCAGCACTATCAGCAGTATAAACACACAGGAGCAGCAAC	696	
DB	610	TCTCATCCGCGAGTGCAGGAGCAACACTATCAGCAGTATACGACAGTGTGTATCAAGT	669	
QY	697	CCAACTTGCAACAACAGGAGCAGATTACAGAAACAGCAAGAGTAGTGTGGTGGGCG	756	
DB	670	CCAGCTTGCAACAGGAGCAGATTACAGAAACCAACAGGAAGTAGTGTGGTGGGCTC	729	
QY	757	ATCATTGCTGCATCATCAAGAGTGNACACAGCTGGAGCAAGTGNACACTGTCTGATTA	816	
DB	730	TTCCCTTGCTCATCATCAAAAGTGAATGCAACTGTACCAAGTAAATATGATGTCTGATTA	789	
QY	817	TGGACAGGCCAAAACCCACACTGAAAATTTCCGAAAAAGTCTTGAGCCAGAAAGCTGCAGA	876	
DB	790	TGGACAGGCCAAAACACACTGCAGCTCCGAAAAAGAACTGGNACCAGAAAGCTGCAGA	849	
QY	877	AGAAAGCTTTGAAAATGGACCAAAAGACTCTCTTCCAGTGAATGCAGCTCCATCCATGTG	936	
DB	850	AGAAAGCTTGGAATGGACCAAAAGAAATCTCTTCCAGTAAATAGCAGCTCCATCCATGTG	909	
QY	937	GACAAGACCACAAATCAAGACTTTAAAGAGAAATTCGGCAGGATGCAGATTCTGTGAT	996	
DB	910	GACAGACCTCAGATCAAAAGACTTTCAAAAGAGAAATTCAGCAGGATGCAGATTCCGTGAT	969	
QY	997	TACAGTACGTCGAGGAGAGTGCACCGTCCGAGTCCCGACTCATGAGGAAGATCATATA	1056	
DB	970	TACAGTGGCCGAGGAGAGTGGTCACTGTTCGAGTACCCACCCATGAAGAGGATCATATA	1029	
QY	1057	CCTATTTTGGAAATTTGCCACAGACAGTTATGACATTTGGGTTTGGGTTTATTTTGAATG	1116	
DB	1030	TCTCTTTTGGAAATTTGCCACAGACAAATTAATGACATTTGGGTTTGGGTTTATTTGAATG	1089	
QY	1117	GACAGACTCTCAAAATGCTGCTGTCAGTGTGCATGTGTCAGTCCAGTCCAGTCCAGGAGA	1176	
DB	1090	GACAGACTCTCCAAACACTGCTGTGAGCGTGCATGTGTCAGTCCAGGATGAGCAGGA	1149	
QY	1177	GGAGGAGGAAGAAATGTCACTTGTGAAGAAAAAGCAAAAAAGCAACCAAGCCCTCT	1236	
DB	1150	GGAGGAGGAAGAAACATCGTGTGTAAGAGAAAGCCAAAGAAATGCCAACAGCCCTTT	1209	
QY	1237	GCTGATGAGATTGTA CTTGTGTACCGGCGGAGCTGTCCAGGAGGAATATATGAGGAG	1296	
DB	1210	GCTGATGAGATTGTCCTGTGTACCGGCGGAGCTGTCTATGAGGAGGTGTATGCTGGCAG	1269	
QY	1297	CCACAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGCATATTCCTACTCTCTGTG	1356	
DB	1270	CCATCAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGCACACTCCTACTCTTTGTG	1329	
QY	1357	GAGGTCCAAAGTCCGTCTACTACAGAGTCTATTATATAGATAGAGTGTGTCTCAAGT	1416	
DB	1330	GGGTCAAAATCAGTCTACTACAGGCTTATATATAGTATAGTAAATGTGTATCAAGT	1389	
QY	1417	CCGAGTCCAGGTTGAGCACAACATGACGTTTAAATTT 1454		
DB	1390	CTGGAGTCTAGGTTGGGAGAGATGACATTTAAATTT 1427		
RESULT 9				
AY150218				
LOCUS	Homo sapiens peripheral benzodiazepine receptor associated protein			
DEFINITION	AY150218 3572 bp mRNA linear PRI 21-APR-2003			
ACCESSION	AY150218 mRNA, complete cds.			
VERSION	AY150218.1 GI:24496472			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			



REFERENCE  
AUTHORS  
TITLE

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 3572)  
Liu,J., Cavalli,L.R., Haddad,B.R. and Papadopoulos,V.  
Molecular cloning, genomic organization, chromosomal mapping and  
subcellular localization of mouse PAP7: a PBR and PKA-R1alpha  
associated protein  
Gene 308, 1-10 (2003)  
2 (bases 1 to 3572)  
Liu,J., Tobin,D., Taaken,K. and Papadopoulos,V.  
Direct Submission  
Submitted (13-SEP-2002) Department of Cell Biology, Georgetown  
Univ./Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20057,  
USA

## FEATURES

## source

## Location/Qualifiers

1..3572  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="placenta"  
27..1613  
/note="PAP7; PBR associated protein"  
/codon\_start=1  
/product="peripheral benzodiazepine receptor associated  
protein"  
protein\_id="AA060219.1"  
/db\_xref="GI:24496473"  
/translation="MAAVLNARLEVSVDLSTDPERRPGAEGAPLLPPPLPPSP  
PSGPGASGQPEGEAAGAAEARRLQRMFGLELYGLALRLFKEKDGKAP  
HPTYEKALVHLKQVLMGPNYPOTCEVGFDFVLGNDRRERMAALNFKSKEDAMVE  
FVKLNRCCHLSTVYASHKIEKEODKRRKEERREERERLQKEERKREER  
ERLRKEERIRIEERLRLQKQOQIMAAINSQVAVQFYAAQVPGNYEQOQIL  
IRLOQEHYQOYMOQLYVLAQQAALQKQEVVAGSLPTSCKVNAVTPSNMPPV  
NQAKTHDSSEKLEPEAAEALENGKESLPVIAAPSMWTRPOIKQFORSDSAGCR  
PRDYSGRVTVRPVTHEEGSYLFWEPATDNCIDIGFVYFEMTDSPTNAVSHVSES  
SDDEEENICEEKKNKPLDDEIVPYRRDCHEEVVAGSHQVPGRGVLLKFP  
DNSYSLWRSKSYIRVYYTR"

## CDS

## ORIGIN

Query Match 72.3%; Score 1055; DB 9; Length 3572;  
Best Local Similarity 87.0%; Pred. No. 1.6e-260;  
Matches 1182; Conservative 0; Mismatches 175; Indels 2; Gaps 2;

QY 97 AGTAAAGATGCGAAAGCCTTTTCATCCAACTTATGAAGAAAACATGAAGTTCTGCGCACT 156  
DB 311 AGAAAAAGATGCGAAAGCATTTTCATCCAACTTATGAAGAAAACATGAAGTTCTGCGCACT 370  
QY 157 GCATAAGCAAGTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGTTGGATTCTT 216  
DB 371 GCATAAGCAAGTCTTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTTGGATTCTT 430  
QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAGATGGGCGAGCTCTGGGAAACATGTCCAAGGA 276  
DB 431 TGATGTGTTGGGAATGACAGAGGAGAGAGATGGGCGAGCCCTGGGAAACATGTCTAAAGA 490  
QY 277 GGATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAGTGTGTGCTCTCTCTCGGCATA 336  
DB 491 GGATGCCATGGTAGAGTTTGTCAAGCTCTTAAATAGTGTGTGCTCTCTCTTCAACATA 550  
QY 337 TGTGTCCTCCACAGAAATGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396  
DB 551 TGTGTCCTCCACAAATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610  
QY 397 GCGAGGCGCGTGAAG 456  
DB 611 GCGAGGCGCGTGAAG 670  
QY 457 GCGAGAGGAG 516  
DB 671 GAG 730  
QY 517 GAGGCTTCGGTGGAAACAGCAAAAGCAGAGATATGGCAGCTTTAAACTCCAGACTGC 576  
DB 731 AAGGCTTCGGTGGAGCAGCAAAAGCAGAGATATGGCAGCTTTAAACTCCAGACTGC 790

QY 577 CGTGCAATTCAGCAGTAGTATGCAGCCAGCAGTATCCAGGGAATACGAAACAACAGCAGAT 636  
DB 791 CGTGCAATTCAGCAGTAGTATGCAGCCAGCAGTATCCAGGGAATACGAAACAACAGCAGAT 850  
QY 637 TCTCATCCGCAAGTTCAGGAGCAGCAGTATCAGCAGTATTAACACACAGCAGCAGCAAC 696  
DB 851 TCTCATCCGCAAGTTCAGGAGCAGCAGTATCAGCAGTATTAACACACAGCAGCAGCAAC 910  
QY 697 CCAACTGCAACAACAGCAGCAGCAGTATCAGAAACAGCAAGAGTAGTATGCTGGGGC 756  
DB 911 CCAGCTTGCAACAACAGCAGCAGCAGTATCAGAAACAGCAAGAGTAGTATGCTGGGGC 970  
QY 757 ATCAATTCGCTGCATCATCAAGGTGAACAACAGCTCGGCAAGTGTATACACTGTCACTTAA 816  
DB 971 TTCTCTTGCCTACATCATCAAAAGTGAATGCAACTGTCAAGTAAATATATGCCAGTTAA 1030  
QY 817 TGGACAGGCCAAACCCACACTGAAATTCGAAAGTCTTGTAGCCAGCAAGAGCTGCAGA 876  
DB 1031 TGGACAGGCCAAACCCACACTGACAGCTCCGAAAGAACTGGAAACCAAGAGCTGCAGA 1090  
QY 877 AGAAGCCTTGGAAATGGACCAAAAGACTCTCTCCAGTGTGTCAGCTCCATCCATG 936  
DB 1091 AGAAGCCTTGGAAATGGACCAAAAGACTCTCTCCAGTGTGTCAGTATAGCAGCTCCATCCATG 1150  
QY 937 GACAAGACCAAAATCAAGACTTTT-AAAGAGAAGATTTCGGCAGGATGCAGATTCTGTGA 995  
DB 1151 GACAAGACCTCAGATCAAGACTTTTCAAAGAGAGATTTCAGCAGGATGCAGATTCCGTGA 1210  
QY 996 TTACAGTAGCTCGAGAGAGTCTGTCAGCTCGAGTCCGAGTCTATAGAGAGAGATCAT 1055  
DB 1211 TTACAGT-GGGCAGGAGAGTGGTCACTGTTCGAGTACCCACCCATGAAGAAGGATCAT 1269  
QY 1056 ACCTATTTGGAAATTCGACAGACAGTATGACATTGGGTTTGGGGTTTATTTGAAT 1115  
DB 1270 ATCTCTTTGGAAATTCGACAGACAAATTTGATGATTGGGTTTGGGGTTTATTTGAAT 1329  
QY 1116 GGACAGACTCTCAAAATGCTGTGTCAGTGTGCATGTTCAGTGCAGTCCAGTGCAGGAGG 1175  
DB 1330 GGACAGACTCTCAAAACACTGTGTGTCAGGTCATGTTCAGTGCAGTCCAGCAGTGCAGG 1389  
QY 1176 AGGAGAGAGAGAGAGTGTCTACTGTGAAGAAAAGCAAAAGAAACGCAACAGCCTC 1235  
DB 1390 AGGAGAGAGAGAGAGAGTGTGTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449  
QY 1236 TGCTCGATGAGATGTCTGTCGTCAGCGGAGTGTCCAGAGAGAGATATGAGGCA 1295  
DB 1450 TGCTCGATGAGATGTGTGTCGTCAGCGGAGTGTCTATGAGGAGGTTGATGTCGCA 1509  
QY 1296 GGCACAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGTATTAATTCCTACTCTCT 1355  
DB 1510 GGCATCAATTCAGGAGAGAGAGTCTATCTCTCAAGTTTGTATTAATTCCTACTCTCT 1569  
QY 1356 GGAGTCCAAAGTTCCTGTACTACAGAGTCTATTATACATAGATAGAGTGTCTGTCAAAG 1415  
DB 1570 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATACATAGATAGAGTGTCTGTCAAAG 1629  
QY 1416 TCCGAGTCCAGGTTGAGCACAACATGACGTTTAAATT 1454  
DB 1630 TCTGGAGTCTAGGTTGGGCGAGAGATGACATTTAAATT 1668

## RESULT 10

## BC034563

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

BC034563 3037 bp mRNA linear PRI 19-NOV-2003  
Homo sapiens acyl-Coenzyme A binding domain containing 3, mRNA  
(cDNA clone IMAGE:3858463), partial cds.

BC034563  
BC034563.2 GI:34785054

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```
QY 1217 AAGAAGCCCAAGCCTCTGCTGGATGAGATTGTACCTGTGTACCGGGGAGCTGTCTCAC 1276
DB 901 AAGAATGCCAACAAGCCTTGTCTGGATGAGATTGTGCTGTGTACCGGGGAGCTGTCTCAT 960
QY 1277 GAGGAAGTATATCCAGGAGCCACAGTATCCAGGAGGGGAGTCTATCTCTCAAGTTT 1336
DB 961 GAGGAGGTGTATCTGCTGGAGCCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTT 1020
QY 1337 GATAATTCCTACTCTCTGTGGAGGTCAGAGTCGCTACTACAGAGTCTATTATATACAG 1396
DB 1021 GACAACTCTACTCTCTGTGGCGGTCAAAATCAGTCTACTACAGAGTCTATTATATACAG 1080
QY 1397 TAGAGTCTGTCTTCCAGGTCCGAGTCCAGGTTCCAGGTTCCAGGTTCCAGGTTTAAATTT 1454
DB 1081 TAAAAATGTTGTACAAAGTCTGAGTCTAGGTTCCGAGGAGGATGACATTAAATTT 1138

RESULT 11
AJ720620
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AJ720620 3492 bp mRNA linear VRT 30-SEP-2004
Gallus gallus mRNA for hypothetical protein, clone 21m6.
AJ720620
ORF1.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1
Caldwell,R.B., Kierzek,A.M., Arakawa,H., Bezzubov,Y., Zaim,J.,
Fiedler,P., Kutter,S., Blagodatski,A., Kostovska,D., Koter,M.,
Plachy,J., Carninci,P., Hayashizaki,Y. and Buerstedde,J.M.
Full-length cDNAs from bursal lymphocytes to facilitate gene
function analysis
Unpublished
2 (bases 1 to 3492)
Caldwell,R.B.
Direct Submission
Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuherberg, GERMANY
Location/Qualifiers
1..3492
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="21m6"
/cell_type="lymphocyte"
/tissue_type="bursa"
/clone_lib="rikeni"
/dev_stage="2 weeks old"
1..3492
/locus_tag="RCJMB04_21m6"
<1..50
/locus_tag="RCJMB04_21m6"
51..1610
/locus_tag="RCJMB04_21m6"
/note="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAG32279.1"
/db_xref="GI:53133900"
/translation="MAAALSARLEVSVDGLTSPNAEAPPCEARPGHSDSAAGRRA
GPGCSQAEAGEEAAALSPERRWGFALFELYLALRPFKEKDGKAPHTYVEKRLV
ALHKVLLGPNPDTCPEVGFDFVLGNDRKEWALGNMSKQEAETEFVKLNRCHL
FSTYVTSKIEKEEKKRREERREERREERLEKKEEKRRREERLEERLEER
RRIEERLMEQKQKQIWAALNSQTAMQFQYAAQYQPNVEQQQILIRIQLEQHQYQ
YMQQYQVQAQQAALQKQEAQVAAAGTPLTTASKVNVPAQGMPSINGQASHD
NPEKELDPEALREALENGPKDSVPVIAAPSMWARPQIKDFKEKIRQDADSVITVGRGE
VTVRVPTHEEGSYLFWFATDSYDIGFVYFEWTDSPNTAVSVHVHVSSESDDEDEE
```

```

NASSBEKAKKNANKPQLDEIVLVYRRDCHEEVYAGSHQVGRGVVLLKFDNSYSLMRS
KTVYRVVYTR"
1611..-3492
/locus_tag="RCJMB04_21m6"

3' UTR
ORIGIN
Query Match 59.3%; Score 865.8; DB 5; Length 3492;
Best Local Similarity 79.5%; Pred. No. 9e-212;
Matches 1039; Conservative 0; Mismatches 262; Indels 6; Gaps 1;

QY 98 GTAAAGATGGCAAGCCCTTTTCATCCAACTTATGAAGAAAACTGAAGTTCTGTGGCACTG 157
DB 303 GAAAAGATGGCAAGCCCTTTTCATCCAACTATGAAGAAAACTCAGACTTGTGGCACTG 362
QY 158 CATAGCAGAGTTCTTTTGGGCCCATATACCCAGACACGTCCTCCCTGAGGTTGAGTTCTTT 217
DB 363 CACAAGCAGTTCTGCTGGGACCTTACAACTCTGATCTTGCCCTGAGTTGGTCTCTT 422
QY 218 GATGTTTGGGAATGATAGGAGGAGAAATGGGAGCTCTGGGAAACATGTCCAAGGAG 277
DB 423 GATGTTGCTGGGGAATGATAGAGGAAGGAATGGCTGCCCTTGGAAACATGTGGAAGCAA 482
QY 278 GATGCCATGTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCGGCATAT 337
DB 483 GAAAGCTATGACAGAAATTTGTTAAGCTCTCTAAATAGTGTCTGCCACTTGTCTCAACGAT 542
QY 338 GTTGGCTCCCAAGATAGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAG 397
DB 543 GTTACTTCTCACAAGATGAAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 602
QY 398 CGAAGCGAGCGTGAAGAGGAAGGAAGCGAGCGCTGCAAAAAGGAAGGAAGGAAGCGGAAG 457
DB 603 CGAAGCGCGCTGAAGAGAGAGAGCTTGAAGCGGTTACAGAAAGGAAGGAAGGAAGGAAG 562
QY 458 CGAAGGAGGAAGAACCGCTGAGCGGAGGAGGAAGAGAGCGCGGATAGAGGAAGAG 517
DB 663 CAAGAAGAAGAGAGAGGCTGAGAAGAGAGGAAGAGGAGGAGGAGGAGGAGGAG 722
QY 518 AGGCTTCGGCTGGACACGACGAAGCAGAGATAATGGGAGCTTTAACTCGCAGAGTCC 577
DB 723 CGGCTTCGGATGGAACAGCAAAAGCAACAGATCATGGGAGCAGCTCAACTCCAGAGTCC 782
QY 578 GTGCAATTCACAGCAGTATGCGAGCCAGCAGATATCCAGGGAACCTACGAAACAAACAGAGAT 637
DB 783 ATGCAATTCACAGCAGTATGCTGCTCAGCAGTATCCGGCACTATGACAGCAGCAGATC 842
QY 638 CTATCCCGCAGCTGCGAGGAGCAGCTATCAGCAGTATTAACACACAGCAGCAGCAACCC 697
DB 843 CTCATTCGACAGCTCAAAGAGCAGCAGCTATCAACAGTATATGCAAGTGTGTATCAAGTC 902
QY 698 CAACTGTCACAAACAGCAGCAGCATTACAGAAACAGCAAGTGGAGCAAGTATGATGATG 751
DB 903 CAGCTTGCACAGCAACAGGCGAGCCTTGAGAAACAGCAGGAGGAGGAGTGTGGCAGCAGCA 962
QY 752 GGGGATCATTTGCTGTCATCAAGGTGAACACAGCTGGAGCAAGTATGATGATGATG 811
DB 963 GGGACACCTCTGACTACTGCTTCCAGGTGAATGTACTTCCCAAGGGGATATGCGTCT 1022
QY 812 GTTAATGGCAGGCGCAAAACCCACACTGAAATTCGAAAAAGTCTTTGAGCAGAGAGT 871
DB 1023 ATTAATGGGCAAGCCAGTGCAACACAGCAACCCCTGAAAAGGAGCTGATCCAGAGGCT 1082
QY 872 GCAGAAAGCCCTTGGAAAAATGGACCAAAAGACTCTCTTCCAGTATTTGACGCTCATCC 931
DB 1083 TTGGAAGAAGCATTTGGAGAATGGACCAAAAGATTTCTGTTCAGTATGATAGTCTCATCG 1142
QY 932 ATGTGCAACAGCAACCAATCAAGACTTTTAAAGAGAGATTTCCGCAAGATGCAGATTC 991
DB 1143 ATGTGGCAGCAGCCAGATTAAGACTTTCAAAGAAAAATCCGCGCAGATGCAGACTCT 1202
QY 992 GTGATTACAGTACGTCGAGGAGAGTCTGCTACCGTCGAGTCCCGACTCATGAGGAAGA 1051
DB 1203 GTGATCACAGTGGGCGGAGGGAGTGTGTACAGTTAGAGTACCACACTCATGAGAGGGG 1262
```



Db 1224 AGGTCTAAATCAGTCTACTACAGAGTCTATTACACTAGATAAAG-----ATGTTA 1273

Qy 1418 CGGAGTCCAGGGTTTGAGCACAAACATGACGTTTAAATTT 1454

Db 1274 CAGAGTTTAGAGTTGGCAGAGAATGCAATTTAAATT 1310

RESULT 13

AX677275 961 bp DNA linear PAT 27-MAR-2003

LOCUS

DEFINITION Sequence 53 from Patent WO02086122.

ACCESSION AX677275

VERSION AX677275.1 GI:29334680

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Legrain, P. and Daviet, L.

TITLE Protein-protein interactions in adipocytes

JOURNAL Patent: WO 02086122-A 53 31-OCT-2002;

Hybrigenics (FR)

FEATURES

source Location/Qualifiers

1..961

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 52.1%; Score 760; DB 6; Length 961;

Best Local Similarity 87.0%; Pred. No. 1.4e-184;

Matches 835; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Qy 439 GGAAGAGAGAGCGGAAGCGAGAGGAGGAGACCGGCTGAGCGGAGGAGAGAGAG 498

Db 1 GGAGGAGAGAGAAACGTAGGAGAGAGAGAGAAAGGCTTCGACGGGAGGAGGAAG 60

Qy 499 GCGCGGATAGAGGAAGAGAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATGGCAGC 558

Db 61 GAGACGATAGAGAGAGAGAGGCTTCGGTTGGACAGCAAAAGCAGCAGATATGGCAGC 120

Qy 559 TTTAAACTCGCAGACTCGGTGCAATTCACGACGATATCGACCCAGCAGTATCCAGGAA 618

Db 121 TTTAAACTCCAGACTCGCGTGCGAGTTCACGACGATATCGACCCCAACAGTATCCAGGAA 180

Qy 619 CTAGCAACACAGCAGATTCATCTCCGCGAGCTCCAGGAGCAGCATTACAGCAGTATAA 678

Db 181 CTACGAAACAGCAGCAAAATTCCTCATCCGCGAGTTCGAGGAGCAACACTATCAGCAGTACAT 240

Qy 679 ACACGAGCAGAGCAAAACCCAACTCGACAAACAGCAGCAGATTACAGAAACAGCAAGA 738

Db 241 GCACGAGTTGATCAAGTCCAGCTTGACAGCAACAGCAGCAATACAGAAACACAGGA 300

Qy 739 AGTAGTGATGGCTGGGGCATCATTTGCCGTGCATCATCAAAGGTGAACACAGCTGGAGCAAG 798

Db 301 AGTAGTAGTGGCTGGGCTCTCTTTGGCTTACATCATCAAAAGTGAATGCAACTGTACCAAG 360

Qy 799 TGATACACTGTCACTTAATGGACAGGCCAAACCCACACTGAAATTCGGAAGAAAGTCCT 858

Db 361 TAATATGATGTCACTTAATGGACAGGCCAAACACACTGACAGCTCCGGAAGAAAGTCT 420

Qy 859 TGACCCAGAGCTCCAGAGAGGAGCTTTGGAAATTTGGACCAAGAGCTCTCTCCAGTAT 918

Db 421 GGAACACAGAGCTCGAGAGAGGAGCCCTGGAGAAATGGACCAAGAAATCTCTTCAGTAA 480

Qy 919 TGCAGCTCCATCCATGTGGACAGACCAACCAAAATCAAAGACTTTTAAAGAGAGAGATTCGGCA 978

Db 481 AGCAGCTCCATCCATGTGGACAGCAGCTCAGATCAAAGACTTCAAAGAGAGATTCAGCA 540

Qy 979 GGATGCAGATTCGTGTGATTAACAGTACGTGAGGAGAGAGTGTGTACCGTCCAGTCCCGAC 1038

Db 541 GGATGCAGATTCGTGTGATTAACAGTGGGCGGAGGAGAGTGTGTCACTGTCGAGTACCCAC 600

Db 241 GAATGCAACTGTACCAAGTAATATGATGTGAGTTAATGGACAGGCCCAAAACACACACTGA 300  
Qy 841 AAATTCGAAAAGTCTTGGAGCAGAACTGAGAAAGGCTTGGAAAATGGACCAA 900  
Db 301 CAGCTCCGAAAAGAACTGGAAACCAAGAGCTGAGAAAGCCCTGGAGAAATGGACCAA 360  
Qy 901 AGACTCTCTTCCAGTGATTTGCAAGTCCATCCATGTTGCAAGACACCAAAATCAAGACTT 960  
Db 361 AGAATCTCTTCCAGTAAATAGCAGCTCCATCCATGTTGCAAGACACCAAAATCAAGACTT 420  
Qy 961 TAAAGAGAAAGATTCGGCAGGATGCAAGTCTGTGATTTACAGTACGTCCAGGAGAAAGTCT 1020  
Db 421 CAAAGAGAAAGATTCAGCAGGATGCAAGTCCGTGATTTACAGTGGCCGAGGAGAGTGT 480  
Qy 1021 CACCGTCCAGTCCGAGCTCATGAGGAGGATCACTTATTTGAATGGACAGACTCTCCAAATGCTGCTGT 1140  
Db 481 CACTGTTCCAGTACCCACCCATGAAGAGGATCATATCTCTTTGGGAATTTGCCACAGA 540  
Qy 1081 CAGTTATGACATTTGGGTTTGGGTTTATTTGAATGGACAGACTCTCCAAATGCTGCTGT 1140  
Db 541 CAATTAATGACATTTGGGTTTGGGTTTATTTGAATGGACAGACTCTCCAAACACTGCTGT 600  
Qy 1141 CAGTGTGCAATGTCAGTGAAGTCCAGTCCAGGAGGAGGAGGAGGAGGAGGAGTCACTTTG 1200  
Db 601 CAGCGTGCAATGTCAGTGAAGTCCAGGATGCAAGGAGGAGGAGGAGGAGGAGGAGT 660  
Qy 1201 TGAAGAAAAGCAAAAGAAAGCCCAAGAGCTCTGCTGGATGAGATTTGATCACTGTGTA 1260  
Db 661 TGAAGAAAAGCAAAAGAAAGTCCCAAGAGCTTTGCTGGATGAGATTTGCTGCTGTGTA 720  
Qy 1261 CCGGGGGAGCTGTCAGGAGGAGTATATGAGGAGGAGGAGGAGGAGGAGGAGT 1320  
Db 721 CCGACGGAGCTGTCTAGGAGGAGTATGCTGGAGGAGGAGGAGGAGGAGGAGT 780  
Qy 1321 CTATCTCTCTCAAGTTTGATAATTTCTTCTCTGAGGAGTCCAAAGTCCGTTCTACTACAG 1380  
Db 781 CTATCTCTCTCAAGTTTGATAATTTCTTCTCTGAGGAGTCCAAAGTCCAAATCAGTCTACTACAG 840  
Qy 1381 AGTCTATTATCTAGATA 1398  
Db 841 AGTCTATTATCTAGATA 858

RESULT 15  
BD242865  
LOCUS  
DEFINITION  
Accession  
BD242865  
VERSION  
BD242865.1  
KEYWORDS  
JP 2002536973-A/16  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2714)  
Valenzuela,D., Yuan,O., Hoffman,H., Hall,J. and Rapiejko,P.  
Secreted proteins and polynucleotides encoding them  
Patent: JP 2002536973-A 16 05-NOV-2002;  
ALPHABET: INC  
OS  
Homo sapiens (human)  
PN  
JP 2002536973-A/16  
PD  
Q5-NOV-2002  
PF  
18-FEB-2000 JP 2000599860  
PR  
18-FEB-1999 US 60/120680,23-APR-1999 US 09/298733 PR  
17-AUG-1999 US 60/149639,23-SEP-1999 US 60/155686 PR  
01-OCT-1999 US 60/157247,29-NOV-1999 US 60/167823 PR  
29-NOV-1999 US 60/167822,15-FEB-2000 US 60/182711 PI DARIO  
VALENZUELA, OLIVE YUAN, HEIDI HOFFMAN, JEFF HALL, PETER PI RAPIEJKO  
PC C12N15/09, A61K38/00, A61P3/10, A61P5/14, A61P11/00, A61P11/06, PC  
A61P19/02,  
PC A61P21/04, A61P25/14, A61P27/02, A61P29/00, A61P31/04, A61P31/10,  
PC A61P31/12,  
PC A61P31/18, A61P31/20, A61P31/22, A61P37/00, A61P37/06, C07K14/435,

PC C12N5/10,  
PC C12P19/34, C12R1/91, C12N15/00, C12N5/00, A61K37/02  
CC Secreted proteins and polynucleotides encoding them FH key  
Location/Qualifiers  
FT source 1..2714  
FT /organism='Homo sapiens (human)'.  
FEATURES  
source  
1..2714  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 43.5%; Score 634.8; DB 6; Length 2714;  
Best Local Similarity 86.2%; Pred. No. 2.8e-152;  
Matches 702; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
Qy 641 ATCCGCCAGCTGCGAGGAGCAGCATTATCAGAGATTAACACCGAGCAGAGCAAAACCCAA 700  
Db 2 ATCCGCCAGTTGCGAGGAGCAACACTATCAGCAGTACATGCGAGCTTGTATCAGTCCAG 61  
Qy 701 CTTGCACAAACACAGGCGCATTACAGAAACAGCAAGAGTAGTGTGCTGGGCATCA 760  
Db 62 CTTGCACAGCAACAGGCGCATTACAGAAACAAACAGGAAGTAGTAGTGGCTGGTTTCC 121  
Qy 761 TTGCTGTATCATTAAGGTGAACACAGCTGGAGCAAGTGTACTGTCTAGTTAATGA 820  
Db 122 TTGCTGTATCATTAAGGTGAACACTGTACCAAGTAATATGATGTCTAGTTAATGA 181  
Qy 821 CAGGCCAAAACCCACACTGAAATTCGAAAAGTCCCTTGAGCCAGAGAGTGCAGAGAA 880  
Db 182 CAGGCCAAAACACACTGACAGCTCCGAAAAGAACTGGAACCAAGAGTGCAGAGAA 241  
Qy 881 GCCTTGGAAAATGGACCAAAAGCTCTCTTCCAGTGATTGCAAGTCCATCCATGTGGACA 940  
Db 242 GCCTTGGAGATGGACCAAAAGAACTCTTCCAGTAATAGCAGCTCCATCCATGTGGACA 301  
Qy 941 AGACCAAAATCAAAAGACTTTAAAGAGAAAGTTCCGGCAGGATGAGATCTGTGATTACA 1000  
Db 302 CGACCTCAGATCAAAAGACTTCAAGAGAAAGATTTCAGCAGGATGCGATTCGTTGATTACA 361  
Qy 1001 GTACGTCGAGGAGAGTGTGTCACCGTCCGAGTCCGAGTCCATGAGGAAGGATCATACCTA 1060  
Db 362 GTGGCCCGAGGAGAGTGTGTCCTGTCGAGTACCACCCATGAAAGAGATCATATCTC 421  
Qy 1061 TTTTGGGAAATTTGCCACAGAGTATTGACATTTGGGTTTGGGTTTATTTTGAATGGACA 1120  
Db 422 TTTTGGGAAATTTGCCACAGACAAATATGACATTTGGGTTTGGGTTTATTTTGAATGGACA 481  
Qy 1121 GACTCTCAAAATGCTGTGTGAGTGTGATGTGTCAGTGTGATGTCAGTGTGAGGAGGAG 1180  
Db 482 GACTCTCAAAACACTGTGTGAGTGTGATGTGTCAGTGTGATGTCAGTGTGAGGAGGAG 541  
Qy 1181 GAGGAAGAAATGTGTCATTTGGAAGAAAGCAAAAGAAAGCGCAAGAGCTCTGCTG 1240  
Db 542 GAAGAAAGAAACATCGGTTGTGAAGAGAAAGCAAAAGAAAGTGCACCAAGCTTTGCTG 601  
Qy 1241 GATGAGATTGATCTGTGTACCGGGGAGCTGTGTCAAGGAGAGTATATGAGGAGCCAC 1300  
Db 602 GATGAGATTGTCCTGTGTACCGAGCGGAGTGTGTATGAGGAGTGTATGTCGAGCCAT 661  
Qy 1301 CAGTATCAGGAGGAGGAGTGTATCTCTCAAGTTTGAATATCTCTCTCTCTGTGGAGG 1360  
Db 662 CAATATCAGGAGGAGGAGTGTATCTCTCAAGTTTGAATATCTCTCTCTCTGTGGAGG 721  
Qy 1361 TCCAGTCCGCTCTACTACAGAGTCTATTATATCTACTAGTAGAGTGTCTGTTTCAAGTCCGG 1420  
Db 722 TCAAAATCAGTCTACTACAGAGTCTATTATATCTACTAGTAAATTTGTTTCAAAAGTCTGG 781  
Qy 1421 AGTCAGGAGTGTGAGCACAACATGAGTTTAAATTT 1454  
Db 782 AGTCAGGAGTGTGAGCAGAAGATGACATTTAAATTT 815

Search completed: August 22, 2005, 22:24:27  
Job time : 6440 secs

---

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 18:02:18 ; Search time 840 Seconds  
(without alignments)  
10282.027 Million cell updates/sec

Title: US-09-762-594-2

Perfect score: 1459

Sequence: 1 gaattcgccgcgcgtcgac.....catgacgtttaatttccttt 1459

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1459	100.0	1459	3	Aaz57038 PBR-assoc
2	1079.6	74.0	2140	13	Ad34285 POSH prot
3	1079.6	74.0	3399	4	Aak52051 Human pol
4	1079.6	74.0	3598	13	Ad34286 POSH prot
5	1079.6	74.0	3598	13	Ad34287 POSH prot
6	1078	73.9	3560	13	Ad34290 POSH prot
7	1073.2	73.6	1481	3	Aaz87207 Human NTA
8	1068.6	73.2	1995	13	Ad34291 POSH prot
9	1055	72.3	3572	13	Ad34289 POSH prot
10	1045.4	71.7	2350	4	Aak53035 Human pol
11	890.4	61.0	3049	13	Ad34288 POSH prot
12	760	52.1	961	8	Acas56933 Human adi
13	634.8	43.5	2714	3	Aaa93116 Human sec
14	293.2	20.1	420	3	Aac77363 Human ORF
15	268.2	18.4	278	4	Aak53762 Murine tr
16	222.4	15.2	231	3	Aac77466 Human ORF
17	211.4	14.5	2450	8	Abz73815 Secured
18	211.4	14.5	2450	10	Abz67399 Human sec
19	204.4	14.0	934	6	Abk24346 DNA encod
20	150.8	10.3	3547	11	Adm02147 Human CDN

21	150	10.3	882	6	ABQ93525	Abq93525 Human CDN
22	142.6	9.8	873	4	Aah72906	Aah72906 Human cer
23	129	8.8	4384	4	Aak77371	Aak77371 Human imm
24	113.4	7.8	527	6	Abk24345	Abk24345 DNA encod
25	110.6	7.6	554	4	Aai17988	Aai17988 Probe #79
26	110.6	7.6	554	4	ABA62955	ABA62955 Human foe
27	110.6	7.6	554	4	Aai42978	Aai42978 Probe #11
28	110.6	7.6	554	4	Aak37154	Aak37154 Human bon
29	110.6	7.6	554	4	Aak11364	Aak11364 Human bra
30	110.6	7.6	554	4	AbS36831	AbS36831 Human liv
31	110.6	7.6	554	6	AbS11145	AbS11145 Human gen
32	110.6	7.6	210204	12	ADQ18927	ADQ18927 Human sof
33	109.2	7.5	306	4	Aai28147	Aai28147 Probe #18
34	109.2	7.5	306	4	ABA76496	ABA76496 Human foe
35	109.2	7.5	306	4	Aai57184	Aai57184 Probe #25
36	109.2	7.5	306	4	Aak51149	Aak51149 Human bon
37	109.2	7.5	306	4	Aak25143	Aak25143 Human bra
38	109.2	7.5	306	4	ABSS0709	ABSS0709 Human liv
39	109.2	7.5	306	6	ABs24667	ABs24667 Human gen
40	100.8	6.9	1495	4	ABL24635	ABL24635 Drosophil
41	92.6	6.3	30191	12	ADQ97654	ADQ97654 Mouse can
42	91.8	6.3	535	5	ABV55196	ABV55196 Human pro
43	88.8	6.1	812	9	ADB83130	ADB83130 Human CDN
44	85.6	5.9	573	4	Aah71226	Aah71226 Human cer
45	84.4	5.8	204803	12	ADQ97348	ADQ97348 Mouse can

ALIGNMENTS

RESULT 1

Aaz57038

ID Aaz57038 standard; cDNA; 1459 BP.

XX

AC Aaz57038;

XX

DT 19-MAY-2000 (first entry)

XX

DE PBR-associated protein (PAP)7 encoding cDNA.

XX

KW Peripheral-type benzodiazepine receptor; PBR; PBR- associated protein;  
KW PAP; cell proliferation; cancer; cell death; cytostatic; neuroprotective;  
KW immunomodulator; antiinfertility; cerebroprotective; atherosclerosis;  
KW Niemann-Pick C; tumour; Alzheimer's disease; developmental disorder;  
KW cholesterol; multiple sclerosis; stress; neurodegenerative disorder;  
KW immune disorder; stroke; PAP7; ss.

XX

OS Mus sp.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX The invention provides isolated peripheral-type benzodiazepine receptor  
CC (PBR) associated proteins (PAPs) and nucleic acids encoding the PAP  
CC proteins. The PAP polynucleotides are a source of primers and probes for  
CC detection, isolation and amplification. PAP ligands or substrates or  
CC antibodies can be labeled and used to detect PAPs, in the diagnosis or  
CC prognosis of disease associated with increased cell proliferation, such  
CC as cancer, or reduced cell death. The diagnostic methods of the invention  
CC can be predictive of diseases involving PBR including gallstones,  
CC atherosclerosis, Niemann-Pick C, Sirtosterolemia, Dystrophy, tumor  
CC proliferation, Snyder's corneal crystalline dystrophy, brain disorders  
CC including Alzheimer's disease, cholesterol metabolism, Tellurium  
CC toxicity, Smith-Lemli-Opitz syndrome, myelinization, developmental  
CC abnormalities, demyelination, Charcot-Marie tooth disease, Pelizaeus-  
CC Merzbacher disease, Multiple sclerosis, and SLA. The methods may also be  
CC useful in prophylactic treatments, or in screening for compounds  
CC effective in prophylactic treatment. The PAPs may be used to identify  
CC inhibitors or activators which allows the identification of drugs or  
CC agents which modulate PBR activity. Inhibitors of PAP may be used in the  
CC treatment or amelioration of conditions such as stress and stroke,  
CC cancer, neurodegenerative disorders, developmental disorders, infertility  
CC and immune disorders. The present sequence represents a cDNA encoding a  
CC PAP7 polypeptide  
XX  
SQ Sequence 1459 BP; 455 A; 298 C; 403 G; 303 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1459; DB 3; Length 1459;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCGGCGCGCTGACCTAAAGTTGAGTTGTTCACTGTAGTCACCGTGTGAAGGT 60  
DB 1 GAATTCGGCGCGCTGACCTAAAGTTGAGTTGTTCACTGTAGTCACCGTGTGAAGGT 60  
QY 61 AGTTTATTTTAAATCAACTTTTCATTGTGCAACTAGTAAAGATGCGAAAGCCTTTCA 120  
DB 61 AGTTTATTTTAAATCAACTTTTCATTGTGCAACTAGTAAAGATGCGAAAGCCTTTCA 120  
QY 121 TCCAACTTATGAAGAAAAAAGTTGAGTTGCGACCTGCATAGCAAGTTCTTTTGGGCCC 180  
DB 121 TCCAACTTATGAAGAAAAAAGTTGAGTTGCGACCTGCATAGCAAGTTCTTTTGGGCCC 180  
QY 181 ATATAACCCAGACACGTCCCTCAGGTTCGATTCTTTTGATGTTCTGGGAATCATAGGAG 240  
DB 181 ATATAACCCAGACACGTCCCTCAGGTTCGATTCTTTTGATGTTCTGGGAATCATAGGAG 240  
QY 241 GAGAGAAATGGCAGCTCTGGGAAACATGTCCAAAGGAGATGCCATGTTAGTTTGTGAA 300  
DB 241 GAGAGAAATGGCAGCTCTGGGAAACATGTCCAAAGGAGATGCCATGTTAGTTTGTGAA 300  
QY 301 GCTTCTTAATAAGTTGTCTCTCTCTCGGCATATGTTGCTGCCACAGAAATAGAGAA 360  
DB 301 GCTTCTTAATAAGTTGTCTCTCTCTCGGCATATGTTGCTGCCACAGAAATAGAGAA 360  
QY 361 GGAAGAAGAGAGAAAGAGGAGGCGGAGGAGGCGAAGGCGAGGAGGAGGAGGAGGAG 420  
DB 361 GGAAGAAGAGAGAAAGAGGAGGCGGAGGAGGCGAAGGCGAGGAGGAGGAGGAGGAG 420  
QY 421 ACCGAGACGGCTCCAAAAGGAGAGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
DB 421 ACCGAGACGGCTCCAAAAGGAGAGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
QY 481 ACCGGAGGAGGAGAGAGGCGGCGATAGAGGAAGAGAGGCTTCGGCTGGAAACAGCAAAA 540  
DB 481 ACCGGAGGAGGAGAGAGGCGGCGATAGAGGAAGAGAGGCTTCGGCTGGAAACAGCAAAA 540  
QY 541 GCAGCAGATATCGCAGCTTTAACTCGCAGACTGCGCTGCAATTCAGCAGTATCGAGC 600  
DB 541 GCAGCAGATATCGCAGCTTTAACTCGCAGACTGCGCTGCAATTCAGCAGTATCGAGC 600  
QY 601 CCAGCAGTATCCAGGGAACTTACGAAACACAGCAGAGATTTCTATCCCGCAGCTGAGGAGCA 660  
DB 601 CCAGCAGTATCCAGGGAACTTACGAAACACAGCAGAGATTTCTATCCCGCAGCTGAGGAGCA 660

QY 661 GCATATCAGCAGTATAAACACACGCGCAGACAAACCCCAACCTGCAACACAGGCAGC 720  
DB 661 GCATATCAGCAGTATAAACACACGCGCAGACAAACCCCAACCTGCAACACAGGCAGC 720  
QY 721 ATTACAGAAACAGCAAGAGTAGTATGCTGGGATCATTCCTGCTCATCATCAAGGT 780  
DB 721 ATTACAGAAACAGCAAGAGTAGTATGCTGGGATCATTCCTGCTCATCATCAAGGT 780  
QY 781 GAACACAGCTGGAGCAAGTGATACACTGTCACTGTATGAGCAGGCGCAACCACTGA 840  
DB 781 GAACACAGCTGGAGCAAGTGATACACTGTCACTGTATGAGCAGGCGCAACCACTGA 840  
QY 841 AAATTCGAAAAAGTCCITTCAGCCAGAGCTCGAGAGAGCTTCGAAAAATGGACCAA 900  
DB 841 AAATTCGAAAAAGTCCITTCAGCCAGAGCTCGAGAGAGCTTCGAAAAATGGACCAA 900  
QY 901 AGACTCTCTTCAGTATGTCAGCTCCATCCATGTGGACAAGACCAACCAATCAAGACTT 960  
DB 901 AGACTCTCTTCAGTATGTCAGCTCCATCCATGTGGACAAGACCAACCAATCAAGACTT 960  
QY 961 TAAAGAGAGATTCCGCGAGATGCGAGATTCCTGTGATTACAGTACGTCGAGGAGAGTCT 1020  
DB 961 TAAAGAGAGATTCCGCGAGATGCGAGATTCCTGTGATTACAGTACGTCGAGGAGAGTCT 1020  
QY 1021 CACCGTCCAGTCCCGACTCATGAGAGAGATCATACCTATTTTGGGAATTTGCCACAGA 1080  
DB 1021 CACCGTCCAGTCCCGACTCATGAGAGAGATCATACCTATTTTGGGAATTTGCCACAGA 1080  
QY 1081 CAGTTATGACATTTGGGTTTGGGTTTATTTTGAATGGACAGACTCTCCAAATGCTGCTGT 1140  
DB 1081 CAGTTATGACATTTGGGTTTGGGTTTATTTTGAATGGACAGACTCTCCAAATGCTGCTGT 1140  
QY 1141 CAGTGTGATGTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
DB 1141 CAGTGTGATGTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
QY 1201 TGAAGAAAAAGCAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1260  
DB 1201 TGAAGAAAAAGCAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1260  
QY 1261 CCGCGGGGAGTGTCCAGGAGAGTATATGACGAGCAGCCAGTATCCAGGAGGAGGAGT 1320  
DB 1261 CCGCGGGGAGTGTCCAGGAGAGTATATGACGAGCAGCCAGTATCCAGGAGGAGGAGT 1320  
QY 1321 CTATCTCTCAAGTTTGAATAATTCCTACTCTCTGTGGAGGTCCAGTCCGTCTACTACAG 1380  
DB 1321 CTATCTCTCAAGTTTGAATAATTCCTACTCTCTGTGGAGGTCCAGTCCGTCTACTACAG 1380  
QY 1381 AGTCTATTATATAGATAGAGTGTGTTTCCAAAGGTCGGAGTCCAGGGTTGAGCACAAC 1440  
DB 1381 AGTCTATTATATAGATAGAGTGTGTTTCCAAAGGTCGGAGTCCAGGGTTGAGCACAAC 1440  
QY 1441 ATGACGTTTAAATTTTCCTTT 1459  
DB 1441 ATGACGTTTAAATTTTCCTTT 1459

RESULT 2  
ADS34285  
ID ADS34285 standard; DNA; 2140 BP.  
XX  
AC ADS34285;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE POSH protein associated DNA #39.  
XX  
KW ds; gene; cytostatic; neurotropic; neuroprotective; antiparkinsonian;  
KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;  
KW ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;

KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.  
OS Homo sapiens.  
XX WO2004078130-A2.  
XX 16-SEP-2004.  
XX 02-MAR-2004; 2004WO-US006308.  
XX 03-MAR-2003; 2003US-0451437P.  
XX 05-MAR-2003; 2003US-0452284P.  
XX 19-MAR-2003; 2003US-0455760P.  
XX 20-MAR-2003; 2003US-0456640P.  
XX 03-APR-2003; 2003US-0460526P.  
XX 04-APR-2003; 2003US-0460792P.  
XX 21-APR-2003; 2003US-0464285P.  
XX 09-MAY-2003; 2003US-0469462P.  
XX 15-MAY-2003; 2003US-0471378P.  
XX 20-MAY-2003; 2003US-0472327P.  
XX 30-MAY-2003; 2003US-0474706P.  
XX 03-JUN-2003; 2003US-0475825P.  
XX 17-JUN-2003; 2003US-0479317P.  
XX 19-JUN-2003; 2003US-0480215P.  
XX 08-AUG-2003; 2003US-0480376P.  
XX 28-AUG-2003; 2003US-0493860P.  
XX 16-SEP-2003; 2003US-0498634P.  
XX 10-NOV-2003; 2003WO-US035712.  
XX 05-FEB-2004; 2004WO-US003600.  
XX 02-MAR-2004; 2004US-0549896P.  
XX (PROT-) PROTEOLOGICS INC.  
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
XX Greener T;  
XX WPI; 2004-662346/64.  
XX Isolated, purified or recombinant complex, useful for identifying an  
XX antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
XX POSH-associated protein (POSH-AP).  
XX Disclosure; SEQ ID NO 49; 374pp; English.  
XX The invention relates to an isolated, purified or recombinant complex (I)  
XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
XX or HRPUD1 and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are  
XX useful for identifying an agent that modulates an activity of a POSH  
XX polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
XX apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
XX of a protein through the secretory pathway, an agent that inhibits the  
XX progression of a neurological disorder, an agent that modulates a POSH  
XX function, an agent that modulates a HRPUD1 function. The methods can be  
XX used for treating a viral infection, for inhibiting an activity of a POSH  
XX -AP in a cell, for treating a POSH-associated disease in a subject. The  
XX POSH-associated disease is viral infection, POSH-associated cancer or  
XX POSH-associated neurological disorder. The methods are useful for  
XX treating or preventing POSH-associated neurological disorder in a subject  
XX e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
XX schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
XX nucleic acid of the invention.  
XX Sequence 2140 BP; 628 A; 461 C; 587 G; 464 T; 0 U; 0 Other;  
Query Match 74.0%; Score 1079.6; DB 13; Length 2140;  
Best Local Similarity 87.2%; Pred. No. 6.1e-281;  
Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;  
XX 97 AGTAAAGATGGCAAGCCCTTCATCCACTATGAGAAAAAAGTGAAGTTCGTGGCACT 156  
||| ||||||||||| ||||||||||||||||||| |||||||

Db 333 AGAAAAAGATGGCAAGCAATTTTCATCAACTATGAGAAAAATGGAAGCTTGTGGCACT 392  
Qy 157 GCATAAGCAAGTTCTTTTGGGGCCCATATAACCCAGACACGCTCCCTCAGGTTGATTTCTT 216  
Db 393 GCATAAGCAAGTTCTTTATGGGGCCCATATAATCAGACACTTGTCTCAGGTTGATTTCTT 452  
Qy 217 TGATGTGTTGGGGAATGATAGGAGGAGAGAAATGGGAGCTCTGGGAAACATGTCCAAGGA 276  
Db 453 TGATGTGTTGGGGAATGACAGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCTTAAGA 512  
Qy 277 GGATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCGGCATA 336  
Db 513 GGATGCCATGGTAGAGTTTGTCAAGCTTCTAAATAAGTGTGTCTCTCTCTCTCAACATA 572  
Qy 337 TGTGTCCTCCACAGAAATAGAGAGGAGAGAAAGAGAGAAAGAGAAAGCGGAGGAGGA 396  
Db 573 TGTGTCCTCCACAAATAGAGAGGAGAGAAAGAGAGAAAGAGAGAAAGAGGAGGAGGA 632  
Qy 397 GCGAAGGCGAGCGTGAAGAGGAGAAACGAGAGCGGCTGCAAAAGGAGAAAGAGAGCGGAA 456  
Db 633 GCGAAGGCGGCTGAGAGGAGAAAGAGAGAAACGCTCTGCAAAAGAGGAGAGAAACGTA 692  
Qy 457 GCGAGAGGAGGAGAAAGACCGGCTGAGACGGGAGGAGAGAGAGCGGCGGATAGAGGAAGA 516  
Db 693 GAGAGAAGAGAGAGGAGAAAGGCTTCGACGGGAGGAGAGAGAGAGAGCGGATAGAGGAAGA 752  
Qy 517 GAGGCTTCGGCTGGAAACGCAAGCAGCAGAGTAATGGCAGCTTTAAACTCGCAGACTGC 576  
Db 753 AAGGCTTCGGCTGGAGCAGCAAAAGCAGCAGAGTAATGGCAGCTTTAAACTCCCACTGC 812  
Qy 577 CGTGCAATTCAGCAGCTATGAGCCAGCAGTATCAGGGAACTACGAAACAAAGCAGAGAT 636  
Db 813 CGTGCAATTCAGCAGCTATGAGCCAGCAGTATCAGGGAACTACGAAACAAAGCAGAGAT 872  
Qy 637 TCTCATCCCGCAGCTGAGGAGCAGCACTATCAGCAGTATAAACAACAGGAGCAAGCAAC 696  
Db 873 TCTCATCCCGCAGCTGAGGAGCAACACTATCAGCAGTATACATGAGCAGTTGTATCAAGT 932  
Qy 697 CCNACCTGCACACACACACGAGCATTACAGAAACAGCAAGAGTAGTAGTGGCTGGGC 756  
Db 933 CCAGCTTGACAGCAACACGAGCATTACAGAAACAAACAGGAGTAGTAGTGGCTGGGC 992  
Qy 757 ATCATTGCTGCTGATCATCAAGAGTGAAACACAGCTGAGCAAGTATACACTGTCAGTTAA 816  
Db 993 TTCTCTGCTGCTGATCATCAAGAGTGAACTGCACTGTCAAGTATATGATGTCAGTTAA 1052  
Qy 817 TGGACAGGCCAAACCCACACTGAAATTCGAAAAAGTCTTTGAGCCAGAACTGCAGA 876  
Db 1053 TGGACAGGCCAAACCCACACTGACAGCTCCGAAAAAGAACTGGAACCCAGAACTGCAGA 1112  
Qy 877 AGAAGCTTGGAAAAATGGACCAAAAGACTCTCTTCAGTGATTTGAGCTCCATCCATGTG 936  
Db 1113 AGAAGCTTGGAAAAATGGACCAAAAGAACTCTCTTCAGTAAATAGCAGCTCCATCCATGTG 1172  
Qy 937 GACAAGCACCACAAATCAAAAGACTTTTAAAGAGAGATTTGCGCAGGATGCGAGTTCTGTGAT 996  
Db 1173 GACAAGCTTCAGTCAAAAGACTTTCAAGAGAAAGATTCAGAGATTCAGATTCCTGTAT 1232  
Qy 997 TACAGTACGTTCGAGGAGAAAGTCTGTCACCGTCCGAGTCCGAGCTCATGAGGAAGATCAT 1056  
Db 1233 TACAGTGGGCGGAGGAGAAAGTGGTCACTGCTTCAGTACCCACCACCATGAAAGAGATCAT 1292  
Qy 1057 CCTATTTTGGGAATTTGCCACACAGAGTTATGACATTTGGGTTTGGGTTTATTTGAATG 1116  
Db 1293 TCTCTTTTGGGAATTTGCCACACAGCAATTTATGACATTTGGGTTTGGGTTTATTTGAATG 1352  
Qy 1117 GACAGACTCTCCAAATGCTGCTGTCAGTGTGATGTGTCAGTGTGTCAGTGTGTCAGGAGGA 1176  
Db 1353 GACAGACTCTCCAAACACTGCTGTGTCAGCTGTCATGTGTCAGTGTGTCAGGAGGATGACACGA 1412  
Qy 1177 GGAGGAGGAGAAAAATGTCACTTTGTGAAGAAAAAGCAAAAAAGAAAGGAAAGGAAAGGCTCT 1236  
Db 1413 GGAGGAGGAGAAAAATGTCGGTTGTGAAGAGAAAGCAAAAAAGAAAGGAAAGGAAAGGCTCT 1472

QY 1237 GCTGGATGAGATTGATACCTGTGTACCGGGGAGCTGTACAGGAGAGATATATGACGGCAG 1296  
 DB 1473 GCTGGATGAGATTGTCCTGTGTACCGGAGCTGTATGAGGAGGTGTATGCTGGCAG 1532  
 QY 1297 CCACCAAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGATATAATTCCTACTCTCTGTG 1356  
 DB 1533 CATCAATATCCAGGAGAGAGTCTATCTCTCAAGTTTGACAACTCTACTCTTTGTG 1592  
 QY 1357 GAGGTCCAAGTCCGCTACTACAGAGTCTATTATATCTAGATAGAGTGTGTTCCAAAGT 1416  
 DB 1593 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATATCTAGATAAAATGTTGTACAAAGT 1652  
 QY 1417 CCGAGTCCAGGTTGAGCACAAATGACGTTTAATTT 1454  
 DB 1653 CTGGAGTCTAGGTTGGGAGAGATGACATTTAATTT 1690

## RESULT 3

AAK52051  
 ID AAK52051 standard; cDNA; 3399 BP.

AC AAK52051;

DT 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 596.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; SB.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.

PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.

PR 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

DR P-PSDB; AAM78918.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful

PT in diagnosis and gene therapy.

XX Claim 1; Page 2141-2144; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 XX  
 SQ Sequence 3399 BP; 1005 A; 640 C; 791 G; 963 T; 0 U; 0 Other;

Query Match 74.0%; Score 1079.6; DB 4; Length 3399;  
 Best Local Similarity 87.2%; Pred. No. 7.7e-281;  
 Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 97 AGTAAAGATGCAAGAGCCTTTCATCCAACTTATGAAGAAAACCTGAAGTTCTGTCGCACT 156  
 DB 297 AGAAAAAGATGCAAGAGCATTTCATCCAACTTATGAAGAAAATTTGAAGTTGTGGCACT 356  
 QY 157 GCATAAGCAAGTTCTTTTGGGCCCATATAAACCACACAGTCCCTGAGGTTGGATTCTT 216  
 DB 357 GCATAAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTGAGTTGGATTCTT 416  
 QY 217 TGATGTGTGGGAATGATAGGAGGAGAGATGGGCGAGCTCTGGGAAAACATCTCCAAGGA 276  
 DB 417 TGATGTGTGGGAATGACAGGAGGAGAGATGGGCGAGCCCTGGGAAAACATGTCTAAAGA 476  
 QY 277 GGATGCCATGTGTAGAGTTTGTGAAGCTTCTAAATAAGTTTGTCTCTCTCTCGGCATA 336  
 DB 477 GGATGCCATGTGTAGAGTTTGTCAAGCTCTTAATAGTTTGTCCATCTCTTTCAACATA 536  
 QY 337 TGTTCGTCCTCCACAGAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396  
 DB 537 TGTTCGTCCTCCACAAATAG 596  
 QY 397 GCGAAGGAGCGTGAAG 456  
 DB 597 GCGAAGGAGCTGTGAAG 556  
 QY 457 GCGAG 516  
 DB 657 GAG 716  
 QY 517 GAGGCTTTCGCTGGAG 576  
 DB 717 AAGGCTTTCGCTGGAG 776  
 QY 577 CGTCAATTCAG 636  
 DB 777 CGTCAATTCAG 836  
 QY 637 TCTCATCCGCGAGCTGAGGAG 696  
 DB 837 TCTCATCCGCGAGTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 896  
 QY 697 CCAACCTGCACAAACAGGAG 756  
 DB 897 CCAGCTTGCACAGCAACAGGAG 956  
 QY 757 ATCATTCCTGCTCATCATCAAGGTGAACAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 816  
 DB 957 TTTCTTCTGCTCATCATCAAGGTGAACAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1016  
 QY 817 TGGACAGGCGAAAAACCCACACTGAAAAATTCGAAAAAAGTCTTTGAGCCAGAGAGCTGCAGA 876  
 DB 1017 TGGACAGGCGAAAAACACACTGACAGCTCCGAAAAAAGTCTGAGAACAGAGAGCTGCAGA 1076  
 QY 877 AGAAGCTTTGGAAAAATGGACCAAAAGAGACTCTCTTCCAGTGATTTGAGCTCCATCTGTCG 936  
 DB 1077 AGAAGCTTTGGAGATGGACCAAAAGAAATCTCTTCCAGTAAATAGCAGCTCCATCTGTCG 1136  
 QY 937 GACAAGACCAAAATCAAGAGCTTTTAAAGAGAGAGATTTCGCGAGGATTCGAGATTCTGTGAT 996  
 DB 1137 GACACACACTCAGATCAAAAGACTTTAAAGAGAGAGATTTCAGCAGGATTCGAGATTCCGTCGAT 1196  
 QY 997 TACAGTACCTCGAGGAGAGAGTGTCTCACCGTCCGAGTCCGAGTCCATCATGAGAGAGATCAT 1056

Db 1197 TACAGTGGCGGAGGAGAGTGGTCACTGTCAGATACCCACCATGAAGAAGATCATA 1256  
Qy 1057 CCTATTGTTGGAAATTTCCACAGACAGATTATGACATTTGGGTTTATTTTGAATG 1116  
Db 1257 TCTCTTTTGGAAATTTGCCACAGACAAATATGACATTTGGGTTGTTATTTTGAATG 1316  
Qy 1117 GACAGACTCTCCAAATGCTGCTGTGACAGTGTGCATGTGATGAGTCCAGTGACGAGGGA 1176  
Db 1317 GACAGACTCTCCAAACACATGCTGTGACAGTGTGCATGTGATGAGTCCAGTGACGAGGGA 1376  
Qy 1177 GGAGGAGGAGAGAAATGTCATTTGTGAAGAAAAAGCAAAAAAGCAACCAAGCTCT 1236  
Db 1377 GGAGGAGGAGAGAAACATCGTTGTGAAGAAAAAGCAAAAAAGCAACCAAGCTCTT 1436  
Qy 1237 GCTGGATGAGATTTGATCTGTACCGCGGAGTGTGACAGAGAAATATATGACGAGCAG 1296  
Db 1437 GCTGGATGAGATTTGCTGTGTACCGCGGAGTGTGACAGAGAAATATATGACGAGCAG 1496  
Qy 1297 CCACAGTATCCAGGAGGGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTGTG 1356  
Db 1497 CCATCAATATCCAGGAGGAGTCTATCTCTCAAGTTTGAATAATCTCTCTCTGTG 1556  
Qy 1357 GAGTCCCAAGTCTCTTACTACAGAGTCTATTACTAGATAGTGTGTTTCCAGGT 1416  
Db 1557 GCGTCAAAATCAGTCTTACTACAGAGTCTATTACTAGATAGTGTGTTTCCAGGT 1616  
Qy 1417 CCGAGTCCAGGTTTGAGCACATGACGCTTTAAATTT 1454  
Db 1617 CTGAGTCTAGGTTTGGCAGAGATGACATTTAAATTT 1654  
  
RESULT 4  
ID ADS34286  
XX ADS34286 standard; DNA; 3598 BP.  
AC ADS34286;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE POSH protein associated DNA #40.  
XX  
KW ds; gene; cytostatic; nontropic; neuroprotective; antiparkinsonian;  
KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;  
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.  
XX  
OS Homo sapiens.  
XX  
PN WO2004078130-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 02-MAR-2004; 2004WO-US006308.  
XX  
PR 03-MAR-2003; 2003US-0451437P.  
PR 05-MAR-2003; 2003US-0452284P.  
PR 19-MAR-2003; 2003US-0455760P.  
PR 20-MAR-2003; 2003US-0456640P.  
PR 03-APR-2003; 2003US-0460526P.  
PR 04-APR-2003; 2003US-0460792P.  
PR 21-APR-2003; 2003US-0464285P.  
PR 09-MAY-2003; 2003US-0469462P.  
PR 15-MAY-2003; 2003US-0471378P.  
PR 20-MAY-2003; 2003US-0472327P.  
PR 30-MAY-2003; 2003US-0474706P.  
PR 03-JUN-2003; 2003US-0475825P.  
PR 17-JUN-2003; 2003US-0479317P.  
PR 19-JUN-2003; 2003US-0480215P.  
PR 19-JUN-2003; 2003US-0480376P.

PR 08-AUG-2003; 2003US-0493860P.  
PR 28-AUG-2003; 2003US-0498634P.  
PR 16-SEP-2003; 2003US-0503931P.  
PR 10-NOV-2003; 2003WO-US035712.  
PR 05-FEB-2004; 2004WO-US003600.  
PR 02-MAR-2004; 2004US-0549896P.  
XX  
PA (PROT-) PROTEOLOGICS INC.  
XX  
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
PI Greener T;  
XX  
XX WPI; 2004-662346/64.  
XX  
PT Isolated, purified or recombinant complex, useful for identifying an  
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
PT POSH-associated protein (POSH-AP).  
XX  
XX Disclosure; SEQ ID NO 50; 374pp; English.  
XX  
CC The invention relates to an isolated, purified or recombinant complex (I)  
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
CC or HERPUD1 and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are  
CC useful for identifying an agent that modulates an activity of a POSH  
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
CC of a protein through the secretory pathway, an agent that inhibits the  
CC progression of a neurological disorder, an agent that modulates a POSH  
CC function, an agent that modulates a HERPUD1 function. The methods can be  
CC used for treating a viral infection, for inhibiting an activity of a POSH  
CC -AP in a cell, for treating a POSH-associated disease in a subject. The  
CC POSH-associated disease is viral infection, POSH-associated cancer or  
CC POSH-associated neurological disorder. The methods are useful for  
CC treating or preventing POSH-associated neurological disorder in a subject  
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
CC nucleic acid of the invention.  
XX  
SQ Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;  
XX  
Query Match 74.0%; Score 1079.6; DB 13; Length 3598;  
Best Local Similarity 87.2%; Pred. No. 7.9e-281;  
Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;  
Qy 97 AGTAAAGATGCAAGGCTTTCATCCAACTATGAAGAAATCTGAAGTTCTGTCGCACT 156  
Db 340 AGAAAGATGCAAGGCTTTCATCCAACTATGAAGAAATCTGAAGTTCTGTCGCACT 399  
Qy 157 GCATAGCAAGTTCTTTTGGGCCCATATATACCCAGACAGTCCCTGAGGTTGGATTCTT 216  
Db 400 GCATAGCAAGTTCTTTTGGGCCCATATATATCCAGACACTTGTCTGAGTTGGATTCTT 459  
Qy 217 TGATGTTGGGGAATGATAGGAGGAGAGATGGGAGCTCTGGGAAACATGTCCAAAGGA 276  
Db 460 TGATGTTGGGGAATGATAGGAGGAGAGATGGGAGCTCTGGGAAACATGTCTAAAGA 519  
Qy 277 GGATGCCATGTAGAGTTTGTGAAGCTTCTTAATAGTGTGTCTCTCTCGCAT 336  
Db 520 GGATGCCATGTAGAGTTTGTGAAGCTTCTTAATAGTGTGTCTCTCTCGCAT 579  
Qy 337 TGTTCGTCCTCCACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396  
Db 580 TGTTCGTCCTCCACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 639  
Qy 397 GCGAAGGCGCTGTAAG 456  
Db 640 GCGAAGGCGCTGTAAG 699  
Qy 457 GCGAGAGGAG 516  
Db 700 GAG 759  
Qy 517 GAGGCTTCGGCTTGGACAGCAAAAGCAGCAGATATATGGCAGCTTTTAACTCGCAGACTGC 576

```
Db 760 AGGCTTCGGTGGACGACGAAAGCGAGATATGGCAGCTTTAACTCCAGACTGC 819
Qy 577 CGTGCAATTCAGCAGTATGCGCCAGCAGTATCCAGGGAACCTACGAACAACAGCAGAT 636
Db 820 CGTGCAATTCAGCAGTATGCGCCAGCAGTATCCAGGGAACCTACGAACAACAGCAGCAAT 879
Qy 637 TCTCATCCGCGAGCTCAGGAGCAGCACTATCGAGTATTAACACACGCGCAGCAAC 696
Db 880 TCTCATCCGCGAGCTCAGGAGCAACACTATCAGCAGTATGCGCAGCTTGTATCAAGT 939
Qy 697 CCAACCTGCACACAACAGCAGCACTTACAGAAACAGCAAGAGTAGTGATGCTGGGC 756
Db 940 CAGCTTGACAGCAACAGCAGCACTTACAGAAACAGCAAGAGTAGTGATGCTGGGC 999
Qy 757 ATCATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCAGTTAA 816
Db 1000 TTCTTTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAAATATGATGTCAAGTAA 1059
Qy 817 TGAAGGCGCAAAACCCACACTGAATAATTCGGAAGAGTCTTTGAGCCAGAGCTGCAGA 876
Db 1060 TGAAGGCGCAAAACCCACACTGACAGCTCCGAAAAAGAACTGGAAACCAAGAGCTGCAGA 1119
Qy 877 AGAAGCCTTGGAAATGGCAACAGCACTCTCTCCAGTGATTGCGAGCTCCATCCATG 936
Db 1120 AGAAGCCTTGGAAATGGCAACAGCACTCTCTCCAGTAAATAGCAGCTCCATCCATG 1179
Qy 937 GACAAGACCAAAATCAAGACTTTAAAGAGAGATTCGGCAGGATGCGAGATCTGTGAT 996
Db 1180 GACAAGCCTCAGATCAAGACTTCAAGAGAGATTCAGCAGGATGCGAGATTCGTGAT 1239
Qy 997 TACAGTACGTGAGGAGAGTGTCTACCGTCCAGTCCCGACTCATGAGAGAGATCATTA 1056
Db 1240 TACAGTGGGCGGAGAGAGTGTCTCTGTCAGTACCACCCATCCATGAAGAGATCATTA 1299
Qy 1057 CCTATTTCGGAAATTCGCACAGACAGTATGACATTGGGTTTGGGGTTTATTTCAATG 1116
Db 1300 TCTCTTTGGAAATTCGCACAGACAGTATGACATTGGGTTTGGGGTTTATTTCAATG 1359
Qy 1117 GACAGACTCTCCAAATGCTGCTGTCAGTGTGCACTGTCAGTGTCCAGTGACAGGAGGA 1176
Db 1360 GACAGACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTGTCCAGCGATGAACGA 1419
Qy 1177 GGAGGAGGAGAAATGTCATCTGTGAAGAAAGCAAAAGAAAGCAAGCCCAACAGCTCT 1236
Db 1420 GGAGGAGGAGAAATGTCATCTGTGAAGAGAAAGCAAAAGAAATGCCAACAAAGCTTT 1479
Qy 1237 GCTGGATGAGATTGTACCTGTGTACCGGGGAGTGTACAGAGAGAGTATATGCGGCAG 1296
Db 1480 GCTGGATGAGATTGTGCTGTGTACCGAGGAGTGTATGAGAGGTTATGCTGGCAG 1539
Qy 1297 CCACAGTATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATAATTCCTACTCTCTGTG 1356
Db 1540 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGACAACTCTACTCTTTGTG 1599
Qy 1357 GAGGTCGAATCCGCTCTACAGAGTCTATTAATACATAGAGTGTGCTGTGTTCCAGGT 1416
Db 1600 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATATAGATAAAATGTTGTACAAAGT 1659
Qy 1417 CCGGAGTCCAGGTTGAGCAGACATGAGCTTTAAATTT 1454
Db 1660 CTGGAGTCTAGGTTGGGCAAGAGATGACATTTAAATTT 1697
```

## RESULT 5

ADS34287

ID ADS34287 standard; DNA; 3598 BP.

XX ADS34287;

AC ADS34287;

XX 02-DEC-2004 (first entry)

XX POSH protein associated DNA #41.

DE

ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian; anticonvulsant; antiviral; neuroleptic; central nervous system; POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1; Ubiquitin ligase; antiviral agent; anti-apoptotic agent; anti-cancer agent; secretory pathway trafficking inhibitor; neurological disorder progression disorder; Alzheimer's disease; Parkinson's disease; Huntington's disease; schizophrenia; Niemann-Pick's disease.

XX Homo sapiens.

XX WO2004078130-A2.

XX 16-SEP-2004.

XX 02-MAR-2004; 2004WO-US006308.

XX 03-MAR-2003; 2003US-0451437P.

XX 05-MAR-2003; 2003US-0452284P.

XX 19-MAR-2003; 2003US-0455760P.

XX 20-MAR-2003; 2003US-0456640P.

XX 03-APR-2003; 2003US-0460526P.

XX 04-APR-2003; 2003US-0460792P.

XX 21-APR-2003; 2003US-0464285P.

XX 09-MAY-2003; 2003US-0469462P.

XX 15-MAY-2003; 2003US-0471378P.

XX 20-MAY-2003; 2003US-0472327P.

XX 30-MAY-2003; 2003US-0474706P.

XX 03-JUN-2003; 2003US-0475825P.

XX 17-JUN-2003; 2003US-0479317P.

XX 19-JUN-2003; 2003US-0480215P.

XX 19-JUN-2003; 2003US-0480376P.

XX 08-AUG-2003; 2003US-0493860P.

XX 28-AUG-2003; 2003US-0498634P.

XX 16-SEP-2003; 2003US-0503931P.

XX 10-NOV-2003; 2003WO-US035712.

XX 05-FEB-2004; 2004WO-US003600.

XX 02-MAR-2004; 2004US-0549896P.

(PROT-) PROTEOLOGICS INC.

XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S; Greener T;

XX WPI; 2004-662346/64.

XX Isolated, purified or recombinant complex, useful for identifying an antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and POSH-associated protein (POSH-AP).

XX Disclosure; SEQ ID NO 51; 374bp; English.

XX The invention relates to an isolated, purified or recombinant complex (I) comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a) or HERPUD1 and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are useful for identifying an agent that modulates an activity of a POSH polypeptide or POSH-AP, for identifying an antiviral agent, an anti-apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking of a protein through the secretory pathway, an agent that inhibits the progression of a neurological disorder, an agent that modulates a POSH function, an agent that modulates a HERPUD1 function. The methods can be used for treating a viral infection, for inhibiting an activity of a POSH-AP in a cell, for treating a POSH-associated disease in a subject. The POSH-associated disease is viral infection, POSH-associated cancer or POSH-associated neurological disorder. The methods are useful for treating or preventing POSH-associated neurological disorder in a subject e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease, schizophrenia, Niemann-Pick's disease. This sequence corresponds to a nucleic acid of the invention.

XX Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;

SQ

Query Match		74.0%; Score 1079.6; DB 13; Length 3598;
Best Local Similarity		87.2%; Pred. No. 7.9e-281;
Matches 1184; Conservative 0; Mismatches 174; Indels 0; Gaps 0;		
Qy	97	AGTAAAGATGGCAAGCTTTTCATCCAACTTATGAAGAAAACCTGAAGTTTCGTGGCACT 156
Db	340	AGAAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAATTTGAAGCTTTGGCACT 399
Qy	157	GCATAAGCAAGTTCTTTTGGCCCATATPAACCCAGACACGTCCTCCTGAGGTTGGATTCTT 216
Db	400	GCATAAGCAAGTTCTTTTGGCCCATATATCCAGACACTTTGTCTCCTGAGGTTGGATTCTT 459
Qy	217	TGATGCTTGGGAATGATAGGAGAGAAATGGCGAGCTCTGGAAACATGTCCAAGGA 276
Db	460	TGATGCTTGGGAATGATAGGAGAGAAATGGCGAGCTCTGGAAACATGTCTAAAGA 519
Qy	277	GGATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAAGTTGTCTCTCTCCGCGATA 336
Db	520	GGATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAAGTTGTCTCTCTCTCCGCGATA 579
Qy	337	TGTTGCGTCCACAGATAGAGAAGGAAGAGAGAAAAGAAAGAAAGCGGAGGAGGA 396
Db	580	TGTTGCGTCCACAAAATAGAGAAGGAAGAGCAAGAAAAGAAAGGAAGGAGGAGGA 639
Qy	397	GCBAAGCAGCGTGAAGAGGAGAAACGAGAGCGGCTCCAAAGGAAGAAAGCGGAA 456
Db	640	GCBAAGCAGCGTGAAGAGGAGAAAGAGAAAGCGTCTGCAAAAGGAGGAGAAACGTAG 699
Qy	457	GCBAAGCAGGAGAGACCGGCTGAGACGCGGAGGAGAGAGAGCGGCGGATAGAGGA 516
Db	700	GAGAGAAGAGAGAAAGGCTTCGACCGGAGGAGAGAGAAAGGAGAGCGGATAGAGGA 759
Qy	517	GAGGCTTCGGTGGAAACAGCAAGACGACAGATAATGGCAGCTTTAACTTCGCAAGTCG 576
Db	760	AAGGCTTCGGTGGAGCAGCAAGACGACAGATAATGGCAGCTTTAACTTCGCAAGTCG 819
Qy	577	CGTGCAATTCAGCAGATGATGAGCCAGCAGATATCCAGGAACTACGAAACAGCAGAT 636
Db	820	CGTGCAATTCAGCAGATGATGAGCCAGCAGATATCCAGGAACTACGAAACAGCAGAAAT 879
Qy	637	TCTCATCCGCGCAGTCGAGGAGCAGCACTATCAGCAGTATAAACACAGCAGGAGCAAC 696
Db	880	TCTCATCCGCGCAGTTGCGAGGAGCACTATCAGCAGTATAATGATGATGATGATCAAGT 939
Qy	697	CCAACCTGCACAAACAGCAGCAGCATTTACAGAAACAGCAAGAGTAGTGCTGGGGC 756
Db	940	CCAGCTTGACAGCAACAGCAGCAGCATTTACAGAAACAGCAAGAGTAGTGCTGGGTC 999
Qy	757	ATCATTGCTGATCATCAAGGTGAACAAGCTGGAGCAAGTGATACACTGTCAGTTAA 816
Db	1000	TTCTGCTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAAGTAA 1059
Qy	817	TGACAGGCGCAAAACCAACACTGAAATTCGAAAGTCTTTGAGCCAGAACTGCAGA 876
Db	1060	TGACAGGCGCAAAACCAACACTGAGCTCCGAAAGAAAGTGAACCAAGAGCTGCAGA 1119
Qy	877	AGAAGCCTTGGAAATGGACAAAGACTCTTTCCAGTGATTGCGAGCTCCATCCATGTG 936
Db	1120	AGAAGCCTTGGAGATGGACCAAAAGATCTTTCCAGTAAATAGCAGCTCCATCCATGTG 1179
Qy	937	GACAAGACCAAAATCAAGAATTTAAAGAGAGATTCGGCAGGATGCAAGTTCTGTGAT 996
Db	1180	GACAAGACCTCAGATCAAAAGACTTCAAGAGAGAGATTCAGCAGGATGCAGATTCCTGTAT 1239
Qy	997	TACAGTACGTCGAGGAGAGTCTGACCGTCCGAGTCCGAGCTCATGAGGAAAGATCAT 1056
Db	1240	TACAGTGGCGGAGGAGAGTGTGCTGCTGTCAGTACCCACCATGAAGAGGATCAT 1299
Qy	1057	CCTATTTTGGGAATTTGCCACAGACAGATGATGACATTTGGGTTTATTTTCAATG 1116
Db	1300	TCTCTTTTGGGAATTTGCCACAGACATGATGACATTTGGGTTTGGGTTGATTTGAATG 1359
Qy	1117	GACAGACTCTCCAAATGCTGCTGTCAGTGTGATGTCAGTGTGATGTCAGTGTGATGTCAGGAGGA 1176
Db	1360	GACAGACTCTCCAAACACTGCTGTGTCAGGTCGATGTGAGTCCAGCGATGACGACGA 1419
Qy	1177	GGAGGAGAGAAAATGTCATCTTGTGAAGAAAAGCAAAAGAACCCCAACAGCTCT 1236
Db	1420	GGAGGAGAGAAAATGTCATCGTTGTGAAGAAAAGCAAAAGAAATGCCAACAGCTTT 1479
Qy	1237	GCTGGATGAGATTTGTACCTGTGTACCGCGGAGCTGTCAAGAGAAATATATGACGGCAG 1296
Db	1480	GCTGGATGAGATTTGTGCTGTGTACCGCGGAGCTGTCAAGAGAAATATATGCTGGCAG 1539
Qy	1297	CCACCATGATCCAGGAGGAGGAGTCTATCTCTCAAGTTCCTCACTCTCTCTCTCTCTGTG 1356
Db	1540	CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTCCTCACTCTCTCTCTCTGTG 1599
Qy	1357	GAGTCCCAAGTCCGCTCTACTACAGAGCTATATATCTAGATAGAGCTGCTGTTCCAAAGT 1416
Db	1600	GCGTCAAAATCAGTCTACTACAGAGCTATATATCTAGATAGAAATGTTGTTTCAAAAGT 1659
Qy	1417	CCGAGTCCAGGTTTGAGCAAAACATGACGCTTTTAAATTT 1454
Db	1660	CTGAGTCTAGGTTGGCAGAGATGACATTTAAATTT 1697
RESULT 6		
ADS34290		
ID	ADS34290 standard; DNA; 3560 BP.	
XX	ADS34290;	
AC	ADS34290;	
XX	02-DEC-2004 (first entry)	
DT	POSH protein associated DNA #44.	
XX	ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;	
XX	anti-convulsant; antiviral; neuroleptic; central nervous system;	
KW	POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI;	
KW	ubiquitin ligase; antiviral agent; anti-apoptotic agent;	
KW	anti-cancer agent; secretory pathway trafficking inhibitor;	
KW	neurological disorder progression disorder; Alzheimer's disease;	
KW	Parkinson's disease; Huntington's disease; schizophrenia;	
KW	Niemann-Pick's disease.	
XX	Homo sapiens.	
OS	WO2004078130-A2.	
XX	16-SEP-2004.	
PD	02-MAR-2004; 2004WO-US0006308.	
PF	03-MAR-2003; 2003US-0451437P.	
XX	05-MAR-2003; 2003US-0452284P.	
PR	19-MAR-2003; 2003US-0455760P.	
PR	20-MAR-2003; 2003US-0456640P.	
PR	03-APR-2003; 2003US-0460526P.	
PR	04-APR-2003; 2003US-0460792P.	
PR	21-APR-2003; 2003US-0464285P.	
PR	09-MAY-2003; 2003US-0469462P.	
PR	15-MAY-2003; 2003US-0471378P.	
PR	20-MAY-2003; 2003US-0472327P.	
PR	30-MAY-2003; 2003US-0474706P.	
PR	03-JUN-2003; 2003US-0475825P.	
PR	17-JUN-2003; 2003US-0479317P.	
PR	19-JUN-2003; 2003US-0480215P.	
PR	19-JUN-2003; 2003US-0480376P.	
PR	08-AUG-2003; 2003US-0493860P.	
PR	28-AUG-2003; 2003US-0498634P.	
PR	16-SEP-2003; 2003US-0503931P.	
PR	10-NOV-2003; 2003WO-US035712.	
PR	05-FEB-2004; 2004WO-US003600.	
PR	02-MAR-2004; 2004US-0549896P.	
XX		



PA (PROT-) PROTEOLOGICS INC.  
XX Taglicht DN, Alroy I, Reies Y, Yaar L, Ben-Avraham D, Tuvia S;  
PI Greener T;  
XX WPI; 2004-662346/64.  
XX Isolated, purified or recombinant complex, useful for identifying an  
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
PT POSH-associated protein (POSH-AP).  
XX Disclosure; SEQ ID NO 54; 374pp; English.  
XX  
XX The invention relates to an isolated, purified or recombinant complex (I)  
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
CC or HRPDI and a ubiquitin ligase (b). Methods using (I), (a) or (b) are  
CC useful for identifying an agent that modulates an activity of a POSH  
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
CC of a protein through the secretory pathway, an agent that inhibits the  
CC progression of a neurological disorder, an agent that modulates a POSH  
CC function, an agent that modulates a HRPDI function. The methods can be  
CC used for treating a viral infection, for inhibiting an activity of a POSH  
CC -AP in a cell, for treating a POSH-associated disease in a subject. The  
CC POSH-associated disease is viral infection, POSH-associated cancer or  
CC POSH-associated neurological disorder. The methods are useful for  
CC treating or preventing POSH-associated neurological disorder in a subject  
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
XX nucleic acid of the invention.  
XX  
XX Sequence 3560 BP; 1059 A; 661 C; 820 G; 1020 T; 0 U; 0 Other;  
Query Match 73.9%; Score 1078; DB 13; Length 3560;  
Best Local Similarity 87.1%; Pred. No. 2.1e-280;  
Matches 1183; Conservative 0; Mismatches 175; Indels 0; Gaps 0;  
97 AGTAAAGATGGCAAGCCCTTCATCACTATGAGAAAGAACTGAGTTCGTGCACCT 156  
318 AGAAAGATGGCAAGCATTTTCATCACTATGAGAAAGAAATTTGAGCTTGTGCACCT 377  
157 GCATAGCAGAGTTCTTTTGGGCCCATATACCCAGACACGTCCTCCCTGAGTTGATTCTT 216  
378 GCATAGCAGAGTTCTTTTGGGCCCATATATTCAGACACTTGTCTGAGTTGGATTCTT 437  
217 TGAATGTTGGGAATGATAGAGGAGAGAAATGGCAGCTCTGGGAAACATGTCGAAGGA 276  
438 TGATGTTGGGAATGACAGGAGGAGAGAAATGGCAGCCCTGGGAAACATGTCGAAGGA 497  
277 GATGCGCATGATAGAGTTTGTGAAGCTTCTAAATAGTGTGTCTCTCTCTCGGCATA 336  
498 GGATGCGCATGTTGGAGTTGTCAAGCTCTTAAATAGGTGTGTCATCTCTTTTCAACATA 557  
337 TGTGTCGTCACAGAAATAGAGAGGAGAGAAAGAGAGAAAGAGAAAGCGGAGAGGA 396  
558 TGTGTCGTCACCAATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 617  
397 GCGAAGGCGCGTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 456  
618 GCGAAGGCGCGTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 677  
457 GCGAAGGAGGAGGA 516  
678 GAGGA 737  
517 GAGGCTTCGCGTGGAAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 576  
738 AAGGCTTCGTTGGAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 797  
577 CGTGCAATTCAGCAGTATGAGCCAGCAGTATCCAGGGAATACGAAACACAGCAGAT 636  
798 CGTGCAATTCAGCAGTATGAGCCAGCAGTATCCAGGGAATACGAAACAGCAGCAAT 857

QY 637 TCTCATCCGCGAGCTGCAGGAGCAGCCTATCAGCAGTATATAAACACAGGAGAGCAAAAC 696  
DB 858 TCTCATCCGCGAGCTGCAGGAGCAGCCTATCAGCAGTATATAAACACAGGAGAGCAAAAC 917  
QY 697 CCAACCTGCACAAACACAGGAGCAGCATTACAGAAAAACAGCAAGAGTAGTAGTGGCTGGGC 756  
DB 918 CCAGCTTGCACAGCAACAGGAGCAGCATTACAGAAAAACAGCAAGAGTAGTAGTGGCTGGGC 977  
QY 757 ATCATTTGCTGATCATCAAGAGTGAAACACAGCTGGAGCAAGTATACACTGTCTGATTTAA 816  
DB 978 TTCTTTGCTGATCATCAAGAGTGAAATGCAATGCAATGCAATGCAATGCAATGCAATG 1037  
QY 817 TCGCAGGCGCAAAACCCACACTGAAATTTCCGAAAAAGTCTTTGAGCCAGAAAGCTGCAGA 876  
DB 1038 TGGCAGGCGCAAAACACACTGACAGCTCCGAAAAAGAACTGGAAACCGGAAGCTGCAGA 1097  
QY 877 AGAAGCCTTGGAAAAATGGACCAAAAGACTCTCTCCAGTAAATAGCAGCTCCATCCATGTG 936  
DB 1098 AGAAGCCTTGGAGATGGACCAAAAGAAATCTCTCCAGTAAATAGCAGCTCCATCCATGTG 1157  
QY 937 GACAGACCAAAATCAAAAGACTTTAAAGAGAAAGATTCCGAGGATGCAAGATTCTGTGAT 996  
DB 1158 GACAGACCTCAGATCAAAAGACTTTCAAGAGAAAGATTTCAGCAGGATGCAAGATTCCGTGAT 1217  
QY 997 TACAGTACGTCGAGGAGAGTCTGTCACGTCGAGTCCGAGCTCATGAGAAAGGATCATA 1056  
DB 1218 TACAGTGGGCGGAGGAGAAAGTGGTCACTGTTGAGTACCCACCCATGAAAGAGGATCATA 1277  
QY 1057 CTTATTTTGGGAATTTGCCACACAGAGATTATGACATGCGTTTGGGGTTTATTTTGAATG 1116  
DB 1278 TCTCTTTTGGGAATTTGCCACACAGCAATATGACATGCGTTTGGGGTGTATTTTGAATG 1337  
QY 1117 GACAGACTCTCAAAATGCTGCTGTCAGTGTGATGTCAGTGTGATGTCAGTGTGATGTCAGGAGGA 1176  
DB 1338 GACAGACTCTCAAAACACTGCTGTGTCAGCGTGTGATGTCAGTGTGATGTCAGGAGGATGACACGA 1397  
QY 1177 GGAGGAGGAGAAATGTCATCTTGTGAAGAAAGCAAAAGAGCAACCAAGCTCT 1236  
DB 1398 GGAGGAGGAGAAAGAAACATCGGTTGTGAAGAAAGCAAAAGAGATGCCAACAGCTTT 1457  
QY 1237 GCTGATGATGATTTGATACCTGTGTATCCGCGGAGCTGTATCCGAGGAGATATATGAGGAG 1296  
DB 1458 GCTGATGATGATTTGCTGCTGTGTACCGAGCGGACTGTATGAGGAGGTTGATGCTGGCAG 1517  
QY 1297 CCACAGATTCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATTTCTACTCTCTG 1356  
DB 1518 CCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGAACACTCTCTCTTTG 1577  
QY 1357 GAGTCCCAAGTCCGCTACTACAGAGTCTATTATATAGATAGAGTGTCTGTTTCCAAAGT 1416  
DB 1578 GCGGTCAAAATCAGTCTACTACAGAGTCTATTATAGATAGATAGATAGATAGATAGAT 1637  
QY 1417 CCGGAGTCCAGGTTTGGAGCAACATGACGTTTAAATTT 1454  
DB 1638 CTGAGTCTAGGTTTGGGAGAGAGATGACATTTAAATTT 1675  
RESULT 7  
AAZ87207  
ID AAZ87207 standard; cDNA; 1481 BP.  
XX  
XX AAZ87207;  
AC  
XX  
XX 08-MAY-2000 (first entry)  
DT  
XX  
XX Human NTAP cDNA clone 998868.  
DE  
XX  
XX Neurotransmission-associated protein; NTAP;  
KW benzodiazepine receptor-associated protein; neurological disease;  
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
KW cerebral neoplasm; multiple sclerosis; drug screening; gene therapy;  
XX antagonist; cancer; AIDS; asthma; Crohn's disease; osteoporosis; ss.



[illegible]

Db 1390 CTGGAGCTAGGTTGGGAGAGATGACATTAAATTT 1427  
RESULT 8  
ADS34291  
ID ADS34291 standard; DNA; 1995 BP.  
XX  
AC ADS34291;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE POSH protein associated DNA #45.  
XX  
ds; gene; cytostatic; neuroprotective; antiparkinsonian;  
KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI1;  
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.  
XX  
OS Homo sapiens.  
XX  
PN WO2004078130-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 02-MAR-2004; 2004WO-US006308.  
XX  
PR 03-MAR-2003; 2003US-0451437P.  
PR 05-MAR-2003; 2003US-0452284P.  
PR 19-MAR-2003; 2003US-0455760P.  
PR 20-MAR-2003; 2003US-0456640P.  
PR 03-APR-2003; 2003US-0460526P.  
PR 04-APR-2003; 2003US-0460792P.  
PR 21-APR-2003; 2003US-0464285P.  
PR 09-MAY-2003; 2003US-0469462P.  
PR 15-MAY-2003; 2003US-0471378P.  
PR 20-MAY-2003; 2003US-0472327P.  
PR 30-MAY-2003; 2003US-0474706P.  
PR 03-JUN-2003; 2003US-0475825P.  
PR 17-JUN-2003; 2003US-0479317P.  
PR 19-JUN-2003; 2003US-0480215P.  
PR 19-JUN-2003; 2003US-0480376P.  
PR 08-AUG-2003; 2003US-0493860P.  
PR 28-AUG-2003; 2003US-0498634P.  
PR 16-SEP-2003; 2003US-0503931P.  
PR 10-NOV-2003; 2003WO-US035712.  
PR 05-FEB-2004; 2004WO-US003600.  
PR 02-MAR-2004; 2004US-0549896P.  
XX  
PA (PROT-) PROTEOLOGICS INC.  
XX  
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
PI Greener T;  
XX  
XX WPI; 2004-662346/64.  
XX  
XX Isolated, purified or recombinant complex, useful for identifying an  
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
PT POSH-associated protein (POSH-AP).  
XX  
XX Disclosure; SEQ ID NO 55; 374pp; English.  
PS  
XX The invention relates to an isolated, purified or recombinant complex (I)  
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
CC or HERPUDI1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are  
CC useful for identifying an agent that modulates an activity of a POSH  
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
CC of a protein through the secretory pathway, an agent that inhibits the

CC progression of a neurological disorder, an agent that modulates a POSH  
CC function, an agent that modulates a HERPUDI1 function. The methods can be  
CC used for treating a viral infection, for inhibiting an activity of a POSH  
CC -AP in a cell, for treating a POSH-associated disease in a subject. The  
CC POSH-associated disease is viral infection, POSH-associated cancer or  
CC POSH-associated neurological disorder. The methods are useful for  
CC treating or preventing POSH-associated neurological disorder in a subject  
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
CC nucleic acid of the invention.  
XX  
SQ Sequence 1995 BP; 628 A; 412 C; 549 G; 406 T; 0 U; 0 Other;  
Query Match 73.2%; Score 1068.6; DB 13; Length 1995;  
Best Local Similarity 87.1%; Pred. No. 5.6e-278;  
Matches 1184; Conservative 0; Mismatches 174; Indels 1; Gaps 1;  
QY 97 ACTAAAAGATGCAAGCCCTTTCATCCAACTTATGAAGAAAACTGAAGTTCGTGGCACT 156  
Db 310 AGAAAAAGATGCAAGCAATTTTCATCCAACTTATGAAGAAAAATGAAGCTTGTGGCACT 369  
QY 157 GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGGTTGGATTCTT 216  
Db 370 GCATAAGCAAGTTCTTTTGGGCCCATATAATCCAGACACTTGTCTCTGAGGTTGGATTCTT 429  
QY 217 TGATGTGTTGGGAATGATAGGAGAGAGAAATGGGCAGCTCTGGGAAACATGTCTCCAAGGA 276  
Db 430 TGATGTGTTGGGAATGACAGAGGAGAGAAATGGGCAGCCCTGGGAAACATGTCTTAAGA 489  
QY 277 GGATGCCATGTTAGAGTTTGTGAAGCTTCTAAATAGTGTGTCTCTCTCTCGGCATA 336  
Db 490 GGATGCCATGTTGGAGTTTGTCAAGCTCTTAAATAGTGTGTCCATCTCTTTTCAACATA 549  
QY 337 TGTTCGCTCCCAAGATAAGAAAGAAAGAG-AGAAAAAGAAAGAGGCGGAGGAGG 395  
Db 550 TGTTCGCTCCCAAAATAGAGAGAGAGCAAGAAAAAAGAAAGAGGAGGAAGAGG 609  
QY 396 AGCGAAGGCGGTGAAGAGGAAGAACGAGAGCGGTGCAAAAGAGAGAGAGAGAGAGCGGA 455  
Db 610 AGCGAAGGCGGTGAAGAGGAAGAGAAACGCTGCAAAAGAGAGAGAGAGAGAAACGTA 669  
QY 456 AGCGAGAGGAGGAAGACCGCTGAGACGGGAGAGAGAGAGGCGCGGATAGAGGAAG 515  
Db 670 GGAGAGAGAGAGAGAGAGAGGCTTCGACGGAGAGAGAGAGAGAGAGAGAGAGAG 729  
QY 516 AGAGGCTTCGGTGGAAACAGCAAGCAAGCAAGTATGCAAGCTTTTAACTCCGACACTG 575  
Db 730 AAAGGCTTCGGTGGAGCAGCAAAAGCAGCAGATAATGTCAGCTTTTAACTCCGACACTG 789  
QY 576 CCGTGCATTCAGCAGATGATGAGCCAGCAGTATCCAGGGAACCTACGAAACACAGCAGA 635  
Db 790 CCGTGCAGTTCAGCAGATGATGAGCCCAACAGTATCCAGGGAACCTACGAAACACAGCAA 849  
QY 636 TTCTCATCCGCGCAGCTGCAGGAGCAGCAGTATCAGCAGTATAAACACACGAGCAGAGCAA 695  
Db 850 TTCTCATCCGCGGTTGCGAGGAGCAACACTATCAGCAGTACATGCGAGAGTTGTATCAAG 909  
QY 696 CCCAACCTGCACAAACAGGAGGAGCATTTACAGAAACAGCAAGAGTAGTGGCTGGGG 755  
Db 910 TCCAGCTTCGACAGCAACAGGAGGAGCATTTACAGAAACAGCAAGAGTAGTGGCTGGGT 969  
QY 756 CATCATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTGTATCACTGTCACTTA 815  
Db 970 CTTCTCTGCTTATCATCATCAAAAGTGAATGCCAACTGTACCAAGTAAATATGATGTCACTTA 1029  
QY 816 ATGGACAGGCCCAAAACCCACACTGCAGTAAATTTCCGAAAAAGTCTTTCAGCCAGAGCTGCAG 875  
Db 1030 ATGGACAGGCCCAAAACACACACTGCAGCTTCGAAAAAGAACTGGAAACAGAGCTGCAG 1089  
QY 876 AAGAAGCCTTGGAAAAATGACCAAGAGACTCTCTTCAGTGAATTCAGCTCCATCCATGAT 935  
Db 1090 AAGAAGCCTTGGAGAAATGACCAAAAGAAATCTCTTCCAGTAAATAGCAGCTCCATCCATGAT 1149

PR	20-MAY-2003; 2003US-0472327P.
PR	30-MAY-2003; 2003US-0474706P.
PR	03-JUN-2003; 2003US-0475825P.
PR	17-JUN-2003; 2003US-0479317P.
PR	19-JUN-2003; 2003US-0480215P.
PR	19-JUN-2003; 2003US-0480376P.
PR	08-AUG-2003; 2003US-0493860P.
PR	28-AUG-2003; 2003US-0498634P.
PR	16-SEP-2003; 2003US-0503931P.
PR	10-NOV-2003; 2003WO-US035712.
PR	05-FEB-2004; 2004WO-US003600.
PR	02-MAR-2004; 2004US-0549896P.
XX	(PROT-) PROTEOLOGICS INC.
PA	
PI	Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
PI	Greener T;
XX	
DR	WPI; 2004-662346/64.
XX	
PT	Isolated, purified or recombinant complex, useful for identifying an
PT	antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
PT	POSH-associated protein (POSH-AP).
XX	
PS	Disclosure; SEQ ID NO 53; 374pp; English.
XX	
CC	The invention relates to an isolated, purified or recombinant complex (I)
CC	comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
CC	or HERPUD1 and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are
CC	useful for identifying an agent that modulates an activity of a POSH
CC	polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
CC	apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
CC	of a protein through the secretory pathway, an agent that inhibits the
CC	progression of a neurological disorder, an agent that modulates a POSH
CC	function, an agent that modulates a HERPUD1 function. The methods can be
CC	used for treating a viral infection, for inhibiting an activity of a POSH
CC	-AP in a cell, for treating a POSH-associated disease in a subject. The
CC	POSH-associated disease is viral infection, POSH-associated cancer or
CC	POSH-associated neurological disorder. The methods are useful for
CC	treating or preventing POSH-associated neurological disorder in a subject
CC	e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC	schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
CC	nucleic acid of the invention.
XX	
SQ	Sequence 3572 BP; 1067 A; 665 C; 826 G; 1013 T; 0 U; 1 Other;
	Query Match            72.3%; Score 1055; DB 13; Length 3572;
	Best Local Similarity   87.0%; Pred. No. 3.5e-274;
	Matches 1182; Conservative   0; Mismatches 175; Indels   2; Gaps   2
QY	97 AGTAAAGATGGCAAGSCCTTTTCATCCAACCTATTGMAAGAAAACTGAAGTTTCGTGCACCT 156
Db	311 AGAAAAGATGGCAAGAACAATTTCATCCAACCTATTGAAGAAAAATTTGAAGCTTTGTGGCACC 370
QY	157 GCATAAGCAAGTCTCTTTTGGGCCCATATATAACCCAGACACGTCCTCGAGTTTGGAATCTTT 216
Db	371 GCATAAGCAAGTCTCTTATGGSCCATATATAATCCAGACACTGTCTCTGAGTTTGGATCTTT 430
QY	217 TGATGTGTTGGGGAATGATAGGAGGAGAGAAATGGCGAGCTCTGGGAAAACATGTCCAAGA 276
Db	431 TGATGTGTTGGGGAATCACAGGAGGAGAGAAATGGCGACGCCCTGGGAAAACATGTCTTAAGA 490
QY	277 GGATGCCATCGTAGAGTTTCTGAGCTTCTAAATAAGTTGTCTCTCTCTCTCGGCATA 336
Db	491 GGATGCCATCGTGGAGTTTGTCAAGCTCTTAAATAGGTTGTGGCCATCTCTTTTCAACATA 550
QY	337 TGTTCCTCCACAGAAATAGAGAAGGAAGAAGAGAAAAAAGAAAGAACGGCGGAGGGA 396
Db	551 TGTTCCTCCACAAATAGAGAAGGAAGGACAGACAAAAAAAGGAGGAGGAGGGA 610
QY	397 GCGAAGCGACGCTGAGAGGAAGAACCGAGAGCGGCTGCAAAAGGAAGAAAGACGCGAA 456
Db	611 GCGAAGCGCGCTGCAAGAGGAAGAAAGAGAGCGCTGCAAAAGGAGGAAGAAACGCTAG 670

Query Match	72.3%	Score 1055;	DB 13;	Length 3572;
Best Local Similarity	87.08;	Pred. No. 3.5e-274;		
Matches 1182;	Conservative 0;	Mismatches 175;	Indels 2;	Gaps 2;
Qy	97	AGTAAAGATGGCAAGCCCTTTCATCCAACTTATGAAGAAAAA	CTGAAGTTGCTGGCACT	156
Db	311	AGAAAAAGATGGCAAGCACTTTCATCCAACTTATGAAGAAAAA	TTGAAGCTTGTGGCACT	370
Qy	157	GCATAGCAAGTCTTTTGGGCCCATATATACCCAGACAGTCCCTCGAGTTGGATCTTT		216
Db	371	GCATAGCAAGTCTTTATGGGCCCATATATATCCAGACACTTGTCTCGAGTTGGATCTTT		430
Qy	217	TGATGTGTTGGGGAATGATAGGAGGAGAGATGGCGAGCTCTGGGAAACATGTCCAAAGA		276
Db	431	TGATGTGTTGGGGAATGCACAGGAGGAGAGATGGCAGCCCTGGGAAACATGTCTAAAGA		490
Qy	277	GGATGCCATGGTAGAGTTTGTGAAGCTTCTTAATAAGTGTCTCTCTCTCGGCATA		336
Db	491	GGATGCCATGGTAGAGTTTGTGAAGCTTCTTAATAAGTGTGTGGCACTCTCTTTTCAACATA		550
Qy	337	TGTTTGCCTCCACAGAAATAGAGAAGGAAGAAGAGAAAAAGAAAGCGCGAGGAGA		396
Db	551	TGTTTGCCTCCACAAATATAGAGAAGGAAGAGCAAGAAAGGAAGGAGGAGAGA		610
Qy	397	GCGAAGCCGCTGAAGAGGAAGAACGAGAGCGGCTCTCAAAAGGAAGAAAGCGGAA		456
Db	611	GCGAAGCGGCGCTGAAGAGGAAGAAAGAGAGCGCTCTCAAAAGGAGGAAGAAACGTAG		670

QY 457 GCGAGAGGAGGAGACCGGCTGAGACGGGAGGAGGAGAGCGCGGATAGAGGA 516  
DB 671 GAGAGAGAGAGAGAGAGGCTTCGACGGAGGAGAGAGAGATACGGATAGAGGA 730  
QY 517 GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCGCAGACTGC 576  
DB 731 AAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCGCAGACTGC 790  
QY 577 CGTGAATTCAGCAGTATGACGCCAGCAGCAGTATCCAGGAACTACGAACAACAGCAGAT 536  
DB 791 CGTGCAGTTCCAGCAGTATGACGCCAGCAGCAGTATCCAGGAACTACGAACAACAGCAGCAAT 850  
QY 637 TCTCATCCCGCAGCTCAGCAGCAGCAGTATCAGCAGTATTAACACACGAGCAGCAAC 696  
DB 851 TCTCATCCCGCAGCTCAGCAGCAGCAGTATCAGCAGTATCAGCAGTATCAGCAGTATCAGT 910  
QY 697 CCAACCTGCACAAACAGCAGCAGTATCAGCAGCAGCAGTATCAGCAGTATCAGCAGTATCAGT 756  
DB 911 CCAGCTTCGACAGCAACAGCAGCAGTATCAGCAGCAGCAGTATCAGCAGTATCAGCAGTATCAGT 970  
QY 757 ATCAATTCGCTGATCATCAAGGTGAACACAGCTGGAGCAAGTATCAGTATCAGTATCAGTATCAGT 816  
DB 971 TTCTCTTCGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGCCAGTTAA 1030  
QY 817 TGGACAGGCGCAAAACCCACACTGAARATTCGGAARAGTCTTTGAGCCAGAGCTGCAGA 876  
DB 1031 TGGACAGGCGCAAAACCCACACTGACAGCTCCGAAAAGAACTGGAACAGAGCTGCAGA 1090  
QY 877 AGAAGCTTTGGAAAATGGACCAAAAGACTCTCTCCAGTGAATTCGAGCTCCATCCATGTC 936  
DB 1091 AGAAGCTTTGGAAAATGGACCAAAAGAACTCTCTCCAGTGAATTCGAGCTCCATCCATGTC 1150  
QY 937 GACAAGACCAAAATCAAGACTTTT-AAAGAGAAGATTCGGCAGGATGAGATTCGTGA 995  
DB 1151 GACAAGACCTCAGATCAAAAGACTTTCAAAAGAGAAGATTCAGCAGGATGAGATTCGTGA 1210  
QY 996 TTACAGTACGTCGAGAGAGTGTCTACCGTCGAGTCCGACTCATGAGGAGGATCAT 1055  
DB 1211 TTACAGT-GGCCAGGAGAGTGTGTCACTGTCGAGTACCCACCCATGAAGAGGATCAT 1269  
QY 1056 ACCTATTTTGGGAAATTTGCCACAGACAGTATGACATTTGGGTTTGGGTTTATTTTGAAT 1115  
DB 1270 ATCTCTTTTGGGAAATTTGCCACAGACAAATGTGACATTTGGGTTTGGGTTTATTTGAAT 1329  
QY 1116 GGAAGACTCTCAAAATGCTGTGTGTCAGTGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1175  
DB 1330 GGAAGACTCTCAAAACACTGCTGTGAGGTTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1389  
QY 1176 AGGAGGAGGAGAAATGTCACTTGTGAAGAAAGCAAAAGAGCGCCCAAGCCCTC 1235  
DB 1390 AGGAGGAGGAGAAACATCGTGTGAAGAGAAAGCCAAAAGAAATGCAACAGCCCTT 1449  
QY 1236 TGCTGATGAGATTGTACCTGTGTACCGCGGAGCTGTCAAGAGGAGTATATGACAGGCA 1295  
DB 1450 TGCTGATGAGATTGTGCTGTGTACCGAGGAGCTGTCAAGAGGAGTATATGATGAGCA 1509  
QY 1296 GCCACAGTATCAGGAGGAGGAGTGTATCTCTCAAGTTTGAATTTCTACTCTCTGT 1355  
DB 1510 GCCATCAATATCAGGAGGAGGAGTGTATCTCTCAAGTTTGAATTTCTACTCTCTGT 1569  
QY 1356 GGAGGTCAGGTCGCTCTACTACAGAGTCTATTTATCTAGATAGAGTCTGTCTTCAAG 1415  
DB 1570 GCGGGTCAAAATCAGTCTACTACAGAGTCTATTTATCTAGATAGAAATTTGTTTCAAG 1629  
QY 1416 TCCGAGTCCAGGGTTGAGCAACATGAGCTTTAAATTT 1454  
DB 1630 TCTGAGTCTAGGGTTGGCAGAGAGATGACATTTAATTT 1668

RESULT 10  
AAK53035  
ID AAK53035 standard; cDNA; 2350 BP.

XX AAK53035;  
XX 06-NOV-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 2564.  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX nervous system disorder; arthritis; inflammation; ss.  
XX Homo sapiens.  
XX WO200157190-A2.  
XX 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US004098.  
XX 03-FEB-2000; 2000US-00496914.  
XX 27-APR-2000; 2000US-00560875.  
XX 20-JUN-2000; 2000US-00598075.  
XX 19-JUL-2000; 2000US-00620325.  
XX 01-SEP-2000; 2000US-00654936.  
XX 15-SEP-2000; 2000US-00663561.  
XX 20-OCT-2000; 2000US-00693325.  
XX 30-NOV-2000; 2000US-00728422.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
XX Ma Y, Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;  
XX Xue AJ, Yang Y, Wehrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
XX P-PSDB; AAM79902.  
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
XX in diagnosis and gene therapy.  
XX Claim 1; Page 4841-4842; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
XX sequence listing were missing at the time of publication  
XX Sequence 2350 BP; 687 A; 497 C; 628 G; 538 T; 0 U; 0 Other;  
SQ  
Query Match 71.7%; Score 1045.4; DB 4; Length 2350;  
Best Local Similarity 86.5%; Pred. No. 1.1e-271;  
Matches 1177; Conservative 0; Mismatches 181; Indels 3; Gaps 2;  
QY 97 AGTAAAGATGGCAAGCCCTTTCATCCAACTTATGAAGAAAACCTGAAGTTGTCGCACT 156  
DB 335 AGAAAAGATGGCAAGCAATTTTCATCCAACTTATGAAGAAAATTTGAAGCTTTGTGCACT 394  
QY 157 GCATAGCAAGTTCCTTTTGGGCCCATATAACCCAGACAGTCCCTCGAGCTTGATTCCT 216  
DB 395 GCATAGCAAGTTCCTTTTGGGCCCATATAATCCAGACATTTGTCCTCGAGTTGATTCCT 454  
QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAACATGTCCAAAGGA 276













Qy	97	AGTAAAGATGCGCAAGCCTTTTCATCCAACCTTATGAAGAAAACTGAAGTTCTGTGGCACT	156
Db	66	AGAAAAAGATGCGCAAGCATTTCATCCAACCTTATGAAGAAAAATTTGAAGCTTTGTGGCACT	125
Qy	157	GCATAAGCAAGCTCTTTTGGGCCCATATAACCCAGACAGTCCCTCGAGTTGGATTCTTT	216
Db	126	GCATAAGCAAGTCTTATGGGCCCATATAATCCAGACACTGTCTCGAGTTGGATTCTTT	185
Qy	217	TGATGTGTGGGGAAATGATAGAGAGAGAATGGGCAGCTCTGGGAAACATGTCCAAGGA	276
Db	186	TGATGTGTGGGGAAATGACAGGAGGAGAGAATGGGCAGCCTCGGAAAAATGCTCTAAAGA	245
Qy	277	GGATGCCATGCTAGAGTTTGTGCAAGCTTCTAAATAAGTGTGCTCTCTCTCTCGGCATA	336
Db	246	GGATGCCATGCTGAGTTTGTGCAAGCTCTTAATATAGGTGTGGCCATCTCTTTTCAACATA	305
Qy	337	TGTTGCGTCCACAGAAATAGAGAAGGAAGAAAGAGAAAAAGAAAGCGCGAGGAGGA	396
Db	306	TGTTGCGTCCACAAAAATAGAGAGGAAGACGAACAAAAAGGAAGGGAAGAGGA	365
Qy	397	GCGAAGGCAGCGTGAAGAGGAAGAACGAGCGGCTGCAAAAGGAAGGAAGAA	450
Db	366	GCGAAGGCGCGTGAAGAGGAAGAAAGAACCTCTGCAAAAGGAGGAAGAA	419

RESULT 15

AAK53762/C

ID AAK53762 standard; cDNA; 278 BP.

XX

AC AAK53762;

XX  
XX  
/ 20 / CCXXV

16-NOV-2001

XX  
T007-A00N-0T IT

XX DE Murjine trans

XX  
DE  
MULTIPLA

XX KW Murine: live

KW Murine; live cell metabol

KW cell metabolism  
KW phospholipid

KW phospholipid replication.

KW replication;  
v v

5000

OS Mus musculus

XX

PN DE20103510-U  
....

XX XX X

PD 07-JUN-2001.

XX

PF 28-FEB-2001;

XX

PR 28-FEB-2001;

XX

PA (LION-) LION

XX XX

DR WPI; 2001-36

XX  
XX  
1  
2  
3  
4  
5  
6  
7  
8  
9

Gene library

Query Match	18.4%	Score 268.2;	DB 4;	Length 278;
Best Local Similarity	98.9%;	Pred. No. 5.3e-62;		
Matches 270;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1187	GAAATGCTCACTTGTGAAGAAAAAGCAAAAGAACGCCAACCAAGCCTCTGCTGGATGAG	1246	
Db	277	GAAATGCTCACTTGTGAAGAAAAAGCAAAAGAACGCCAACCAAGCCTCTGCTGGATGAG	218	
Qy	1247	ATTGTACTCTGTACCGCGGGGACTGTCAAGGAAAGTATATGACGAGCACCCACCAAGTAT	1306	
Db	217	ATTGTACTCTGTACCGCGGGGACTGTCAAGGAAAGTATATGACGAGCACCCACCAAGTAT	158	
Qy	1307	CCAGGGAGGGAGTCTATCTCCTCAAGTTTGATAATTCCTACTCTCTGTGGAGGTCCAAG	1366	
Db	157	CCAGGGAGGGAGTCTATCTCCTCAAGTTTGATAATTCCTACTCTCTGTGGAGGTCCAAG	98	
Qy	1367	TCCGTCTACTACAGAGTCTATTATCTAGATAGAGCTGCTGTTCCAAAGGTCGGAGTCCA	1426	
Db	97	TCCGTCTACTACAGAGTCTATTATCTAGATAGAGCTGCTGTTCCAGGGTCGGAGTCTA	38	
Qy	1427	GGGTTGAGCAACAATGACGGTTTAAATTCCTTT	1459	
Db	37	GGGTTGAGCAACAATGACGGTTTAAATTCCTTT	5	

Search completed: August 22, 2005, 22:38:35

Job time : 848 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 18:19:23 ; Search time 5114 Seconds  
(without alignments)

10859.556 Million cell updates/sec

Title: US-09-762-594-2

Perfect score: 1459

Sequence: 1 gaattcgccgcgcgtcgac.....catgaactttaatttccttt 1459

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1337.4	91.7	3433	3 AK030371	Mus muscu
2	1124.4	77.1	1301	9 AY403085	Mus muscu
3	1078	73.9	3560	3 BC045533	Homo sapi
4	1068.6	73.2	1995	3 BC039575	Homo sapi
5	927.2	63.6	1301	9 AY403083	Homo sapi
6	917	62.9	1301	9 AY403084	Pan trogl
7	686.6	47.1	726	6 BY757076	BY757076
8	670	45.9	884	5 BU153320	ACENCOURT
9	635.4	43.6	653	7 CN702497	E0463B05-
10	632.8	43.4	653	7 CF534335	UI-M-FY0-
11	613.8	42.1	822	7 CK600219	ACENCOURT
12	597.4	40.9	757	4 BG294067	ACENCOURT
13	585.4	40.1	603	7 CK627880	ACENCOURT
14	576.6	39.5	1087	1 AL546667	ip04804.y
15	576.4	39.5	792	7 CN310226	ACENCOURT
16	574.2	39.4	742	7 CK960730	ACENCOURT
17	574.2	39.4	761	7 CO555047	ACENCOURT
18	555.6	38.1	932	5 BQ223172	ACENCOURT
19	551.4	37.8	569	7 CK618123	ml05h04.y
20	524	35.9	543	6 CA557620	K0230C03-
21	522.8	35.8	765	5 BU941643	ACENCOURT
22	516.2	35.4	887	5 BU121963	ACENCOURT
23	515.4	35.3	533	8 CC183474	XB462 Bay
24	508.4	34.8	670	7 CK979508	4110992 B

25	499.8	34.3	1108	4 BG254119	BG254119
26	498.4	34.2	516	7 CF175140	602367007
27	495.4	34.0	513	7 CF173967	B0949D03-
28	491.2	33.7	737	4 BG776473	CF173967
29	490.8	33.6	773	5 BU361227	B0931E10-
30	487.6	33.4	707	4 BG705952	602663560
31	485.6	33.3	493	2 BB704602	B0361227
32	483.8	33.2	618	7 CR540238	BG705952
33	482	33.0	755	1 AJ743784	602669271
34	481.2	33.0	696	7 CN532805	BB704602
35	481.2	33.0	760	7 CK638629	CR540238
36	473.2	32.4	486	6 CA558293	DKFZP459L
37	465.8	31.9	581	5 BP293357	BB704602
38	464.2	31.8	768	7 CF741325	BB704602
39	462.4	31.7	480	6 CA557699	CF741325
40	456.4	31.3	776	5 BU243679	UI-M-GH0-
41	452.4	31.0	600	4 BI990454	CF741325
42	452.2	31.0	625	2 AM963720	K0231D08-
43	445.4	30.5	649	2 BE916087	BU243679
44	443.8	30.4	918	7 CF240143	BI990454
45	438	30.0	589	6 CB455265	4071-42 M

#### ALIGNMENTS

RESULT 1	AK030371	AK030371	3433 bp	mrna	linear	HTC 03-APR-2004
LOCUS	Mus musculus	Mus musculus	adult male pituitary gland cDNA, RIKEN full-length			
DEFINITION	enriched library, clone:530404D19 product:GOLGI RESIDENT PROTEIN	enriched library, clone:530404D19 product:GOLGI RESIDENT PROTEIN	GPCP60 homolog [Homo sapiens], full insert sequence.			
ACCESSION	AK030371	AK030371.1	GI:26326368			
VERSION	HTC; CAP trapper.					
KEYWORDS	Mus musculus (house mouse)					
SOURCE	Mus musculus					
ORGANISM	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mus.					
REFERENCE	1	Carninci, P., Itoh, M., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
AUTHORS	High-efficiency full-length cDNA cloning					
TITLE	Normalizaton and subtraction of cap-trapper-selected cDNAs to					
JOURNAL	prepare full-length cDNA libraries for rapid discovery of new genes					
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)					
PUBMED	20499374					
REFERENCE	11042159					
AUTHORS	3					
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Izawa, M., Ohara, E., Watanabe, K., Yoneda, Y., Ishikawa, T., Tanaka, T., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.					
AUTHORS	RIKEN integrated sequence analysis (RISA) system-384-format					
TITLE	sequencing pipeline with 384 multicapillary sequencer					
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)					
MEDLINE	20530913					
PUBMED	11076861					
REFERENCE	4					
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the					
TITLE	FANTOM Consortium.					
JOURNAL	Functional annotation of a full-length mouse cDNA collection					
MEDLINE	Nature 409, 685-690 (2001)					
PUBMED	11076861					
REFERENCE	5					
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research					

Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 3433)  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, F., Iotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.

**TITLE**  
**JOURNAL**  
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.jp,  
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)

**COMMENT**  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.jp/  
 URL: http://fantom.gsc.riken.jp/.

# **FEATURES**

## **source**

Location/Qualifiers  
 1..3433  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:5330404D19"  
 /db\_xref="taxon:10090"  
 /clone="5330404D19"  
 /sex="male"  
 /tissue\_type="pituitary gland"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="adult"  
 61..1638

## **CDS**

/note="unnamed protein product; GOLGI RESIDENT PROTEIN  
 GCP60 homolog [Homo sapiens] (SPTR|Q9H3P7, evidence:  
 FASTY, 88.8%ID, 100%length, match=1575)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAC26928.1"  
 /db\_xref="GI:26326369"  
 /translation="MAQLNVQLVSLDGLTSLPDSERPCGAGAPPOTPPSPAPGN  
 GLGSGSGQORPGPAAAGAAEAREMHQWGLRGLYGLALFYKIDKGAHPHT  
 YEKUKFVALHQLVLLGPNPDTSPVGFVLDNDRREWAALGNMSEKEDAMVEFK  
 LNKCCPLLSAIVASHRIEKEERKAEERKREERERLQKEERKREKEDR  
 LRREERERIEERLRLFOQKQIQAALNSQTAVFQOYAAQYPGNVYEQOILIRQ  
 LQEHYQYQWQIYQVLAQQAALQKQEVVMAGASLPASSKVNTPAGASDLSVNGQ  
 AKHTNSSEKVLPEAAEALENGPKDSLPIAAGPMWTRPOIKDFKEKIRODASVI  
 TVRGEVTVRVPTHEEGSLFWEPATDSYDIFGVYFEWTDSPNAVSVHYSESDSE  
 EEEBENVTCEKAKNANKPLIDELIVPVYRRDCHVEVYAGSHQYVGRGVYLLKFDNS  
 YSLWRSKSVYVYRTR"  
 3416..3421

## **polyA\_signal**

3416..3421

## **polyA\_site**

3433

## **ORIGIN**

Query Match 91.7%; Score 1337.4; DB 3; Length 3433;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 1347; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 97 AGTAAAGATGTCACAAAGCCTTTTCATCAACTTATGAAGAAAACTGAAGTTCTGTCGCACT 156  
 DB 336 AATAAAGATGTCACAAAGCCTTTTCATCAACTTATGAAGAAAACTGAAGTTCTGTCGCACT 395  
 QY 157 GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGGTTGGATTCTT 216  
 DB 396 GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGGTTGGATTCTT 455  
 QY 217 TGATGTGTTGGGAATGATAGAGAGAGAAATGGGAGCTCTGGGAAAACATGTCACAGGA 276  
 DB 456 TGATGTGTTGGGAATGATAGAGAGAGAAATGGGAGCTCTGGGAAAACATGTCACAGGA 515  
 QY 277 GCATGCCATGGTAGAGTTGTGAAGCTTCTAATAAGTTGTTCTCTCTCTCTCTCTCTCGCAT 336  
 DB 516 GGATGCCATGGTAGAGTTGTGAAGCTTCTAATAAGTTGTTCTCTCTCTCTCTCTCGCAT 575  
 QY 337 TGTTGCGTCCCAAGAAATAGAGGAAGAAAGAAAGAAAGAAAGCGGAGGAGGA 396  
 DB 576 TGTTGCGTCCCAAGAAATAGAGGAAGAAAGAAAGAAAGAAAGCGGAGGAGGA 635  
 QY 397 GCGAAGGCAAGCTGAAGAGAAAGAACGAGAGCGGCTGCAAAAGAGAAAGAAAGCGGAA 456  
 DB 636 GCGAAGGCAAGCTGAAGAGAAAGAACGAGAGCGGCTGCAAAAGAGAAAGAAAGCGGAA 695  
 QY 457 GCGAGAGGAGGAAGACCGCTGAGACGCGGAGGAGGAGGCGCGGATAGAGGAAGA 516  
 DB 696 GCGAGAGGAGGAAGACCGCTGAGACGCGGAGGAGGAGGCGCGGATAGAGGAAGA 755  
 QY 517 GAGGCTTTCGGCTGGAAACAGCAAAAGCAGCAGATAATGGCAGCTTTTAAACTCCGCACTGC 576  
 DB 756 GAGGCTTTCGGCTGGAAACAGCAAAAGCAGCAGATAATGGCAGCTTTTAAACTCCGCACTGC 815  
 QY 577 CGTGCAATTCAGCAGATATGACGCCAGCAGATATCAGGGAACTACGAAACAACAGCAGAT 636  
 DB 816 CGTGCAATTCAGCAGATATGACGCCAGCAGATATCAGGGAACTACGAAACAACAGCAGAT 875  
 QY 637 TCTCATCCGCGAGCTGACGAGCAGCAGTATCAGCAGATATAAACACGAGGAGAGCAAAAC 696  
 DB 876 TCTCATCCGCGAGCTGACGAGCAGCAGTATCAGCAGATATAAACACGAGGAGAGCAAAAC 935  
 QY 697 CCAACCTGCACAAACACGAGCAGCATTACAGAAACAGCAAGAGTAGTGATGGCTGGGCG 756  
 DB 936 CCAGCTTGCACAAACACGAGCAGCATTACAGAAACAGCAAGAGTAGTGATGGCTGGGCG 995  
 QY 757 ATCATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACATCTGATTTAA 816  
 DB 996 ATCATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACATCTGATTTAA 1055  
 QY 817 TGGACAGGCCAAACCCACACTGAAATTCGAAAAAGTCCCTTGAGCCAGAAAGCTGCAGA 876  
 DB 1056 TGGACAGGCCAAACCCACACTGAAATTCGAAAAAGTCCCTTGAGCCAGAAAGCTGCAGA 1115  
 QY 877 AGAAGCTTTGAAAAATGGACCAAAAGAGCTCTCTTCAGTGAATTCAGAGCTCCATCCATGTG 936  
 DB 1116 AGAAGCTTTGAAAAATGGACCAAAAGAGCTCTCTTCAGTGAATTCAGAGCTCCATCCATGTG 1175  
 QY 937 GACAGACCAAAATCAAGAATTAAGAGAAAGATTGGCAGGATGCAAGATCTGTGAT 996  
 DB 1176 GACAGACCAAAATCAAGAATTAAGAGAAAGATTGGCAGGATGCAAGATCTGTGAT 1235  
 QY 997 TACAGTAGTCGAGGAGAGTTCGTCACCGTCGAGTCCCGACTCATGAGAAAGGATCAT 1056  
 DB 1236 TACAGTAGTCGAGGAGAGTTCGTCACCGTCGAGTCCCGACTCATGAGAAAGGATCAT 1295  
 QY 1057 CCTATTTTGGGAATTTGCCACAGACAGATTATGACATTTGGGTTTATTTTGAATG 1116  
 DB 1296 CCTATTTTGGGAATTTGCCACAGACAGATTATGACATTTGGGTTTATTTTGAATG 1355  
 QY 1117 GACAGACTCTCAAAATGCTGCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1176  
 DB 1356 GACAGACTCTCAAAATGCTGCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1415  
 QY 1177 GGAGGAGGAAGAAATGTCTACCTTTGTGAAGAAAAAGCAAAAGAAAGCGCCAAAGCTCT 1236

```
Db      1416  GGAGGAGGAGAAATGTCACTGTGAAGAAAAAGCAAAAGAAAGCCCAACAGCCTCT 1475
Qy      1237  GCTGGATGAGATTGTACCTGTGTACCGGGGGAAGTGTACAGAGAAAGTATATGCGGCGAG 1296
Db      1476  GCTGGATGAGATTGTACCTGTGTACCGGGGGAAGTGTACAGAGAAAGTATATGCGGCGAG 1535
Qy      1297  CCACCATGATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATTAATCTCTCTGTG 1356
Db      1536  CCACCATGATCCAGGAGGAGGAGTCTATCTCTCAAGTTTGATTAATCTCTCTGTG 1595
Qy      1357  GAGGTCCAAAGTCGCTACTACAGAGTCTATTATACAGAGTCTGTGTTCCAGGT 1416
Db      1596  GAGGTCCAAAGTCGCTACTACAGAGTCTATTATACAGAGTCTGTGTTCCAGGT 1655
Qy      1417  CCGAGTCCAGGTTTCAGACACATGACGCTTTAAATTTCTTT 1459
Db      1656  CCGAGTCTAGGTTTCAGACACATGACGCTTTAAATTTCTTT 1698

RESULT 2
AY403085
LOCUS      1301 bp      DNA      linear      GSS 15-DEC-2003
DEFINITION Mus musculus GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION  AY403085
VERSION    AY403085.1
KEYWORDS   GI:39759068
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

REFERENCE  1 (bases 1 to 1301)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Adams,M.D. and Cargill,M.
TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE  2 (bases 1 to 1301)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE     Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
            source          1..1301
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
            gene            <1..>1301
                        /gene="GOCAP1"
                        /locus_tag="HCM1439"
            ORIGIN
Query Match      77.1%; Score 1124.4; DB 9; Length 1301;
Best Local Similarity 86.9%; Pred. No. 1.5e-277;
Matches 1131; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy      99  TAAAGATGGCAAGCCTTTTCATCCAACTTATGAAGAAAAAAGTGAAGTTCTGGCACTGC 158
Db      1  TAAAGATGGCAAGCCTTTTCATCCAACTTATGAAGAAAAAAGTGAAGTTCTGGCACTGC 60
Qy      159  ATAAAGCAAGTTCTTTGGGCGCCATATAACCCAGACAGTCCCTGAGGTTGGATCTTTG 218
Db      61  ATAAGCAAGTTCTTTGGGCGCCATATAACCCAGACAGTCCCTGAGGTTGGATCTTTG 120
```

```
Qy      219  ATGTGTTGGGGAATGATAGGAGGAGAAATGGGCGAGCTCTGGGAAACATGTCTCAAGGAGG 278
Db      121  ATGTGTTGGGGAATGATAGGAGGAGAAATGGGCGAGCTCTGGGAAACATGTCTCAAGGAGG 180
Qy      279  ATGCCATGCTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCGGCATATG 338
Db      181  ATGCCATGCTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCGGCATATG 240
Qy      339  TTGCGTCCCAAGATAGAGAAAGGAGAAAGAGAGAAAAAGAAAGGCGGAGGAGGAGC 398
Db      241  TTGCGTCCCAAGATAGAGAAAGGAGAAAGAGAGAAAAAGAAAGNNNNNNNNNNNNNN 300
Qy      399  GAAGGAGCGCTCAAGAGGAGAAAGAGAGCGGCTGCAAAAGAGAAAGAGAGAGCGGAGC 458
Db      301  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
Qy      459  GAGAGGAGGAGAACCGGCTGAGACGGGAGGAGAGAGAGCGGCGGATAGAGGAAGAGA 518
Db      361  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420
Qy      519  GGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATATGGCAGCTTTAAACTCGCAGCTGCCG 578
Db      421  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy      579  TGCATATTCAGCAGTATGAGCCCGCAGCTATCCAGGGAACTACGAAACACAGCAGATTC 638
Db      481  TGCATATTCAGCAGTATGAGCCCGCAGCTATCCAGGGAACTACGAAACACAGCAGATTC 540
Qy      639  TCATCGCCAGCTGCGAGGAGCAGCTATACAGCAGTATATAACACAGGCGAGCAACCC 698
Db      541  TCATCGCCAGCTGCGAGGAGCAGCTATACAGCAGTATATGAGCAGTATATATCAAGTCC 600
Qy      699  AACCTGTCACAAACAGGCGAGCATTTACAGAAAACAGCAAGTAGTGTGGTGGGGCAT 758
Db      601  AGCTTGCACAAACAGGCGAGCATTTACAGAAAACAGCAAGTAGTGTGGTGGGGCAT 660
Qy      759  CATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACATGTCTAGTTAATG 818
Db      661  CATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACATGTCTAGTTAATG 720
Qy      819  GACAGGCCAAAACCCACACTGAAATTCGAAAGATTCCTGAGCCAGAGTGTGCAAGAG 878
Db      721  GACAGGCCAAAACCCACACTGAAATTCGAAAGATTCCTGAGCCAGAGTGTGCAAGAG 780
Qy      879  AAGCCTTGGAAAATGGACCAAAAGACTCTCTTCCAGTGAATTCAGCTCCATCCATCTGGA 938
Db      781  AAGCCTTGGAAAATGGACCAAAAGACTCTCTTCCAGTGAATTCAGCTCCATCCATCTGGA 840
Qy      939  CAAGACCAAAATCAAAAGCTTTTAAAGAGAAAGATTCGGCAGGATGCAGATTCGTGATTA 998
Db      841  CAAGACCAAAATCAAAAGCTTTTAAAGAGAAAGATTCGGCAGGATGCAGATTCGTGATTA 900
Qy      999  CAGTACGTGCGAGGAGAGTCTGTCACGCTCCGAGTCCGACTCATGAGGAAGATCATACC 1058
Db      901  CAGTACGTGCGAGGAGAGTCTGTCACGCTCCGAGTCCGACTCATGAGGAAGATCATACC 960
Qy      1059  TATTTTGGGAAATTCACACAGACAGTTATGACATTTGGGTTTGGGGTTTATTTTGAATGA 1118
Db      961  TATTTTGGGAAATTCACACAGACAGTTATGACATTTGGGTTTGGGGTTTATTTTGAATGA 1020
Qy      1119  CAGACTCTCCAAATGCTGTGTGTCAGTGCATGTCTCAGTGCAGTCCAGTGCAGGAGGAGG 1178
Db      1021  CAGACTCTCCAAATGCTGTGTGTCAGTGCATGTCTCAGTGCAGTCCAGTGCAGGAGGAGG 1080
Qy      1179  AGGAGGAGAAAAATGTCTCTGTGAAGAAAAAGCAAAAGAAAGCGCAACAGACCTCTGC 1238
Db      1081  AGGAGGAGAAAAATGTCTCTGTGAAGAAAAAGCAAAAGAAAGCGCAACAGACCTCTGC 1140
Qy      1239  TCGATGAGATTCCTACTGTGTACCGCGGAGTGTCTCAGGAGAGTATATGAGGCGAGCC 1298
Db      1141  TCGATGAGATTCCTACTGTGTACCGCGGAGTGTCTCAGGAGAGTATATGAGGCGAGCC 1200
Qy      1299  ACCAGTATCCAGGAGGGGAGTCTATCTCTCAAGTTTGTATTAATTCCTACTCTCTGTGGA 1358
```

```
Db      1201  ACCAGTATCCAGGAGGGAGTCTATCTCTCCTCAGTTTGATTAATTCCTACTCTCTGTGA 1260
Qy      1359  GGTCCAAGTCCTCTACTACAGAGTCTATTATACTAGATAG 1399
Db      1261  GGTCCAAGTCCTCTACTACAGAGTCTATTATACTAGATAG 1301

RESULT 3
BC045533 3560 bp mRNA linear HTC 04-MAR-2003
LOCUS clone IMAGE:5259930, mRNA.
DEFINITION BC045533
ACCESSION BC045533
VERSION BC045533.1 GI:28374435
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3560)
Strausberg,R.
Direct Submission
Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaobs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGR1) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systembiology.org
contact: amadan@systembiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 106 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15826851
This clone has the following problem: retained intron.
Location/Qualifiers
1. 3560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5259930"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

FEATURES
source
Query Match 73.9%; Score 1078; DB 3; Length 3560;
Best Local Similarity 87.1%; Pred. No. 1.7e-265;
Matches 1183; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

Qy      97  AGTAAAGATGGCAAGCCCTTCATCCAACTTATGAAGAAAACTGAAGTTCGTGGCACT 156
Db      318  AGAAAAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAATTCGAAGCTTGTGGCACT 377
Qy      157  GCATAGCAAGTCTTTTGGGCCCATATAACCCAGACAGTCCCTCGAGGTTGGATTCTT 216
Db      378  GCATAAGCAAGTCTTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTTGGATTCTT 437
Qy      217  TGATGTGTTGGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAAAATGTCCAAGGA 276
Db      438  TGATGTGTTGGGGAATGACAGGAGGAGAGATGGGCAGCCCTGGGNAACATGTCTAAAGA 497
```

```
Qy      277  GGATCCCATGTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCTCGGCATA 336
Db      498  GGATGCCATGTGGAGTTTGTCAAGCTCTTAAATAGGTGTGGCCATCTCTCTTTCAACATA 557
Qy      337  TGTTCGTGCCACAGATAGAGAAAGAAAGAGAGAAAAAGAAAGGCGGAGAGGAGA 396
Db      558  TGTTCGTGCCACAAATATAGAGAAAGAGAGAAAGAAAAAGAAAGGAGAGAGAGA 617
Qy      397  GCGAAGGCGGTGAAGAGAGAAAGACGAGAGCGGCTGCAAAAGGAAAGAGAAAGCGGAA 456
Db      618  GCGAAGGCGGCGTGAAGAGAGAAAGAGAGAAAGAGAAAGGAGAGAGAAAGCGTAG 677
Qy      457  GCGAGAGAGAGAGAGAGCGGCTGAGAGCGGAGAGAGAGAGAGAGCGGCGGATAGAGAGA 516
Db      678  GAGAGAAAGAGAGAGAGAGGCTTCGACGGGAGAGAGAGAGAGAGAGAGAGAGAGA 737
Qy      517  GAGGCTTCGGCTGGAAACAGCAAAAGCAGCAGATATATGCGAGCTTTAAACTCGCAGACTGC 576
Db      738  AAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATATGCGAGCTTTAAACTCCAGACTGC 797
Qy      577  CGTGCATTCGAGCAGTATGCGCCAGCAGATATCCAGGGAACTACGAAACAAACAGCAGAT 636
Db      798  CGTGCAGTTCGAGCAGTATGCGCCAAACAGTATCCAGGGAACTACGAAACAAACAGCAGCAAT 857
Qy      637  TCTCATCGCCAGCTGCGAGGAGCAGCAGTATCAGCAGTATTAACACACGAGGAGAGCAAAAC 696
Db      858  TCTCATCGCCAGTTCGAGGAGCAACACTATCAGCAGTATACATGCGCAGGTTGTATCAAGT 917
Qy      697  CCAACCTGCACAAACAAAGCAGGAGCAGATACAGAAACAGCAAGAGTAGTGTGGCTGGGCG 756
Db      918  CAGCTTCGACAGCAAGCAGGAGCAGTATACAGAAACAAACAGAGAGTAGTGTGGCTGGGTC 977
Qy      757  ATCATTCGCTGCATCATCAAAAGTGAAACACAGCTGGAGCAAGTAGTATACATGTCTGATTAA 816
Db      978  TTCTCTGCTACATCATCAAAAGTGAAATGCAACTGTACCAAGTATATATGATGTCTGATTAA 1037
Qy      817  TGGACAGGCGCAAAACCCACACTGAAATTCGAAAAAGTCTTGGAGCAGAGAGCTGCAGA 876
Db      1038  TGGACAGGCGCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAAACCGGAAAGCTGCAGA 1097
Qy      877  AGAAGCTTGGAAAAATGACCAAAAGACTCTCTTCAGTGAATGCGAGCTCCATCCATGTG 936
Db      1098  AGAAGCTTGGAGATGACCCAAAGAGATCTCTTCAGTATATAGAGCTCCATCCATGTG 1157
Qy      937  GACAAAGCAAAATCAAAGACTTTTAAAGAGAAAGATTCGGCAGGATGAGATTCGTGTGAT 996
Db      1158  GACACGACCTCAGATCAAAGACTTCAAAGAGAAAGATTCAGCAGGATGAGATTCGCTGAT 1217
Qy      997  TACAGTACGTGAGGAGAGAGTCCGTACCGTCCGAGTCCGACTCATGAGGAGGATCATATA 1056
Db      1218  TACAGTGGGCGGAGGAGAGAGTGGTGCATGTCGAGTACCCACCCCAAGAGAGATCATATA 1277
Qy      1057  CTTATTTTGGGAAATTTGCCACAGACAGATTATGACATGGGTTTGGGGTTTATTTTGAATG 1116
Db      1278  TCTCTTTTGGGAAATTTGCCACAGACATATATGACATTTGGGTTTGGGGTGTATTTTGAATG 1337
Qy      1117  GACAGACTCTCCAAATGCTGCTGTCAGTGTGTCATGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 1176
Db      1338  GACAGACTCTCCAAACACTGCTGTGTCAGCGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 1397
Qy      1177  GGAGAGAGAGAAATGTCACTTGTGAAGAAAAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAG 1236
Db      1398  GGAGAGAGAGAGAAAAATCGGTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
Qy      1237  GCTGATGAGATTGTGTCCTGTGTACCGCGGAGCTGTGTCAGAGGAGAGATATATGAGGAGAG 1296
Db      1458  GCTGATGAGATTGTGCTGTGTGTCACCGAGGAGCTGTGTCAGGAGGAGTGTGTCGTCGAGAG 1517
Qy      1297  CCACAGTATCCAGGAGGAGAGTCTATCTCTCAAGTTTGTGATTAATCTCTCTCTCTCTCTG 1356
Db      1518  CCATCAATATCCAGGAGGAGAGTCTATCTCTCAAGTTTGTGATTAATCTCTCTCTCTCTCTG 1577
```





Db 1570 GCGGTCAAAATCAGTCTCTACAGAGCTATTATCTAGATAAAATGTTGTACAAG 1629

Qy 1416 TCCGGAGTCAGGGTTGAGCACACATGACGTTTAAATT 1454

Db 1630 TCTGGAGCTAGGGTTGGGAGAAGATGCATTTAAATT 1668

RESULT 5	
AY403083	
LOCUS	AY403083 1301 bp DNA linear GSS 15-DEC-2003
DEFINITION	Homo sapiens GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

Query Match	63.6%	Score	927.2;	DB	9;	Length	1301;
Best Local Similarity	77.5%	Pred. No.	7.2e-227;				
Matches 1007; Conservative		0; Mismatches	232;				
						Indels	0; Gaps
							0;

Qy	160	TAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGGTTGGATCTTTGA	219
Db	62	TAAGCAAGTTCTTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTTGGATCTTTGA	121

Qy 280 TGCCATGCTAGAGTTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCGGCATATGT 339

Db 182 TGCCATGCTGGAGTTTGTCAAGCTCTTAAATAGGTGTGGCCATCTCTTTTCAACATATGT 241

Qy		400	AAGCAGCGTGAAAGGGAAGAACGAGAGCGGGCTGCAAAAGGAAGAAAGAGAACGCCGGAAGCG	459
Dd		302	NNN	361
Qy		460	AGAGGAGGAAGA CCGGCTGAGACGGGNGGAGGAAGAGAGCGCGCGGTATAGAGAAAGAGAG	519
Dd		362	NNN	421
Qy		520	GCTTCGGCTGGAA CAGCAAAAGCAGCAGATAAATGGCAGCTTTAAAACTCGCAGACTCCCGT	579
Dd		422	NNN	481
Qy		580	GCAATTCAGAGTAGTGCAGCCAGCAGATTACAGGGAACTAGCAAACAACAGCAGCATTTCT	639
Dd		482	GCAATTCAGAGTAGTGCAGCCAAAGTAGTACCAGGGAACTAAGAAAGCAGCAGCAAAATTTCT	541
Qy		640	CATCCGCCAGCTGCAGGAGCAGCACATATCAGCAGTAGTAACAACAGCAGCAGCAACCCA	699
Dd		542	CATCCGCCAGTTGGAGAGCAACACTATCAGCAGTAGCATGCAGCAGTTGTATCAAGTCCA	601
Qy		700	ACCTGCAACAACAAAGCAGCAGCATTTACAGAAACAGCAAGAGTAGTAGTGGCTGGGGCATC	759
Dd		602	GCTTGCAACAACAGCAGCAGCATTTACAGAAACAACAGGAAGTAGTAGTGGCTGGGCTTC	661
Qy		760	ATTGCCCTGCATCATCAAAAGTGAAACHACAGCTGAGGCAAGTGATACACTGTCAAGTTAATGG	819
Dd		662	CTTGCCCTACATCATCAAAAGTGAACTGTAACCAAGTAATATGATGTCAAGTTAATGG	721
Qy		820	ACAGGCCAAAACCCACACTCAAAATTTCCGAAAAGTCCTTGAGCCAGAAAGCTCGCAAGA	879
Dd		722	ACAGGCCAAAACACACACTGACAGCTCCGNAANAAGAACTGGAACACAGAGCTGCAGAGA	781
Qy		880	AGCCTTGGAAAAATGGACCAAAAGACTCTCTTCAGTGATTTGCAGCTCCATCCATGTGGAC	939
Dd		782	AGCCTTGGGAAATGGACCAAAAGAACTCTCTTCAGTAATAGCAGCTCCATCCATGTGGAC	841
Qy		940	AAGACCACAATCAAGAAGCTTTAAAGAGAGATTTCCGCAGGATGCAGATCTCTGTGATTAC	999
Dd		842	ACAGACTCAGATCAAGAAGCTTTCAAGAGAAAGATTACAGCAGGATGCAGATTTCCGTGATTAC	901
Qy		1000	AGTACGTCGAGGAGAAGTCGTCAACCGTCCCAGTCCCCAGCTCATGAGGAAAGATCATACCT	1059
Dd		902	AGTGGGCGAGGAGAAGTTGTCACCTGTCAGTACCCACCCATGAAGAGGATCATATCT	961
Qy		1060	ATTTTGGGAATTTGCCACAGACAGTTATAGCATTTGGGTTTTGGGTTTTTTGAATGGAC	1119
Dd		962	CTTTTGGGAAATTTGCCACAGACAATATGACATTTGGGTTTGGGTTGATTTTGAATGGAC	1021
Qy		1120	AGACTCTCCAAATGCTGCTGTCTGAGTGCAATGTGCAGTGAGTCCAGTGACGAGGAGGA	1179
Dd		1022	AGACTCTCCAAACACTGCTGTCTGAGCGTGCAATGTCTGAGTCCAGCGATGACACAGAGGA	1081
Qy		1180	GGAGGAGAAAATGTCACTTTGTGAAGAAAAGCAAAAAAGAACCCAAACAGGCTCTGTCT	1239
Dd		1082	GSAAGAAGAAAA CATCGTTGTCAAGAGAAAGCAAAAAAGAA TGCCAA CAAGCCTTTGCT	1141
Qy		1240	GGATGAGATTGTACTGTGTACCGCGGGACTGTCTACGAGGAAGTATATGCAGGCAGCCA	1299
Dd		1142	GSATGAGATTGTGCTGTGTGTACCGACGGGACTGTCTATGAGGAGGTGTATGCTGGCAGCCA	1201
Qy		1300	CCAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGATTAATTCCTACTCTCTGTGGAG	1359
Dd		1202	TCAATATCCAGGAGGAGTCTATCTCTCAAGTTTGACAACTCCCTACTCTCTTTGGCG	1261
Qy		1360	GTCCAAGTCGGTCTACTACAGAGTCATTATACTAGATA	1398
Dd		1262	GTCAAAATCAGTCTACTACAGAGTCTATTATACTAGATA	1300

RESULT 6
AY403084
LOCUS AY403084 1301 bp DNA linear GSS 15-DEC-2003
DEFINITION Pan troglodytes GCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,





JOURNAL  
COMMENT

Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-i@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13203 row: k column: 02  
 High quality sequence stop: 645.  
 Location/Qualifiers  
 1. .884  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6012409"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 92"  
 /notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 2.5 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 45.9%; Score 670; DB 5; Length 884;  
 Best Local Similarity 85.8%; Pred. No. 8.2e-161;  
 Matches 756; Conservative 0; Mismatches 126; Indels 1; Gaps 1;

QY 464 GAGGAAGACCGGCTGAGACGGGAGGAGGAGGAGCGGCGGATAGAGGAAGAGCGCTT 523  
 DB 2 GRAGAGGAAGGCTTCGACGGGAGGAGGAGGAAGAGAGCGGATAGAGGAAGAGCGCTT 61

QY 524 CGGCTGGAAACAGCAAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCAA 593  
 DB 62 CGGTTGGAGCAGCAAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCGAG 121

QY 584 TTCCACAGTATCCAGCCCGCAGTATCCAGGGAATCAGGAACACGACGATCTTCATC 643  
 DB 122 TTCACAGTATCCAGCCCGCAGTATCCAGGGAATCAGGAACACGACGATCTTCATC 181

QY 644 CGCAGCTGCAGGACGACACTATCAGCAGTATAAACCAGGACGACGAAACCCAACT 703  
 DB 182 CGCAGTTCGAGAGCAACACTATCAGCAGTATCAGCAGGATGTATCAAGTCCAGCTT 241

QY 704 GCACAAACACGCGCAGCATTTACAGAAACAGCAAGAAGTAGTAGTGCTGGGGCATATTG 763  
 DB 242 GCACAGCAACGCGCAGCATTTACAGAAACACAGCAAGAAGTAGTAGTGCTGGGCTTCCTTG 301

QY 764 CCTGCATCATCAAGGTGAACACAGCTGGAGCAGGTATACACTGTCAGTTAATGACAG 823  
 DB 302 CCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTAGTTAATGACAG 361

QY 824 GCCAAACCCACACTCAAAATCCGAAAAGTCTTTGAGCCAGAACCTCGAGAAGGCC 883  
 DB 362 GCCAAACACACTGACAGCTCCGAAAAGAACTGGAACCAAGCTCGAGAAGGCC 421

QY 884 TTGGAAAATGGACCAAAAGACTCTCTTCCAGTGATTCGAGCTCCATTCATGTGGACAAGA 943  
 DB 422 CTGGAGATGGACCAAAAGAACTCTCTCCAGTAATAGCAGCTCCATTCATGTGGACACGA 481

QY 944 CCACAAATCAAGACTTTAAAGAGAAGATTGGCAGGATGCAGATTCTGTGATTACAGTA 1003  
 DB 482 CCTCAGATCAAGACTTTCAAGAGAAGATTTCAGCAGGATGCAGATTCCGTGATTACAGTG 541

QY 1004 CGTCGAGGAGAGTCTGCTCCCGTCCGAGTCCGAGCTCATGAGGAAGGATCATACCTATT 1063  
 DB 542 GGCAGGAGGAGAGTGTCTACTGTTTCAGTACCCACCACCATCAAGGAAGGATCATATCTCTT 601

QY 1064 TGGGAATTTGCCACAGACAGTATATGACATTGGGTTTGGGGTTTATTTTGAATGGACAGAC 1123

Db 602 TGGGAATTTGCCACAGACAAATTATGACATTGGGTTTGGGTTTGAATGCACAGAC 661  
 QY 1124 TCTCCAAATGCTGCTGCTGAGTGTGCATGTGCAGTGAGTCCAGTCCAGGAGGAGGAGGAG 1183  
 Db 662 TCTCCAAACACTGCTGCTGAGCGTGCATGTGCAGTGAGTCCAGCGATGACGACGAGGAGAA 721  
 QY 1184 GAAGAAAATGTCTACTTGTGAAGAAAAGCAAAAAGAACGCCAACCAAGCTCTGCTGGAT 1243  
 Db 722 GAAGANAACATCGTTGTGAAGAGAAAGCCCAAAAGATGGCAACCAAGCTTTGCTGGAT 781  
 QY 1244 GAGATTGTACCTGTGTACCGCGGAGACTGTCTCAGGAGAAAGTATATGACGAGCAGCCACCAG 1303  
 Db 782 GAGATTGTGCTGTGTACCCGAGGACTGTCTATGGAGGAGTGTATGCTGCGACCCATCAA 841  
 QY 1304 TAT-CCAGGAGGAGGAGTCTATCTCTCAAGTTTGAATAATTC 1345  
 Db 842 TATCCAGGAGGAGAGAGTCTATCTCTCCTCAGTTTGACCACTCC 884

## RESULT 9

CN702497  
 LOCUS E0463B05-5 NIA Mouse E11.5 whole embryo cDNA library (Long) Mus  
 DEFINITION musculus cDNA clone NIA:E0463B05 IMAGE:30874576 5', mRNA sequence.  
 CN702497  
 ACCESSION CN702497.1 GI:47471246  
 VERSION  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 653)  
 AUTHORS Sharov A.A., Piao Y., Matoba R., Dudekula D.B., Qian Y.,  
 Vanuren V., Falco G., Martin P.R., Stagg C.A., Basse J.C.,  
 Wang Y., Carter M.G., Hamatani T., Aiba K., Akutsu H., Sharova L.,  
 Tanaka T.S., Kimbrell W.L., Yoshikawa T., Jaradat S.A., Pantano S.,  
 Nagaraja R., Boheler K.R., Taub D., Hodes R.J., Longo D.L.,  
 Schlessinger D., Kellier J., Klotz E., Kelsae G., Umezawa A.,  
 Vescevi A.L., Rosant J., Kunath T., Hogan B.L., Curci A.,  
 D'Urso M., Kelsae J., Hide W. and Ko M.S.  
 Transcriptional analysis of mouse stem cells and early embryos  
 PLoS Biol. 1 (3), 410-419 (2003)  
 Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@lgsun.grc.nia.nih.gov  
 Plate: E0463 row: B column: 05  
 Seq primer: M13 Reverse  
 High quality sequence stop: 653  
 POLYA=No.

TITLE  
JOURNAL  
COMMENTFEATURES  
source

1. .653  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="niaEST:E0463B05-5"  
 /db\_xref="taxon:10090"  
 /clone="NIA:E0463B05 IMAGE:30874576"  
 /tissue\_type="whole embryo including extraembryonic  
 tissues" at 11.5-days postcoitum"  
 /dev\_stage="E11.5"  
 /lab\_host="DH10B"  
 /clone\_lib="NIA Mouse E11.5 whole embryo cDNA library  
 (Long)"  
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;  
 Site 2: NotI; Mouse cDNA project by the Laboratory of  
 Genetics, National Institute on Aging (NIA), Intramural  
 Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cdna>).  
 This is a long-transcript enriched cDNA library (Ref.  
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total  
 RNAs were extracted from a pool of 3 embryos at 11.5-days







Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 757)

# REFERENCE AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM10372 row: k column: 08

High quality sequence stop: 680.

Location/Qualifiers

## FEATURES

Source

1..757  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4502911"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_94"  
/note="Organ: eye; Vector: pCMV-Sport6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 40.9%; Score 597.4; DB 4; Length 757;  
Best Local Similarity 96.6%; Pred. No. 3.6e-142;  
Matches 679; Conservative 0; Mismatches 6; Indels 18; Gaps 6;

QY 758 TCATTGCTTCATCAAGGTGACACAGCTGGAGGAGTACATGTCAGTTAAT 817

Db 1 TCATTGCTTCATCAAGGTGACACAGCTGGAGGAGTACATGTCAGTTAAT 60

QY 818 GGACAGGCCAAAACCCACACTGAAATTCGAAAAAGTCTTTGAGCGAGAAGCTGCAGAA 877

Db 61 GGACAGGCCAAAACCCACACTGAAATTCGAAAAAGTCTTTGAGCGAGAAGCTGCAGAA 120

QY 878 GAAGCCTTGGAAAATCGACAAAGACTCTCTTCAGTGATTCAGCTCCATCCATGTGG 937

Db 121 GAAGCCTTGGAAAATCGACAAAGACTCTCTTCAGTGATTCAGCTCCATCCATGTGG 180

QY 938 ACAGACCAAAATCAAGACTTTAAGAGAGATTCGGCAGGATCCAGATTCGTGTATT 997

Db 181 ACAAGACCAAAATCAAGACTTTAAGAGAGATTCGGCAGGATCCAGATTCGTGTATT 240

QY 998 ACAGTACGTCGAGGAGAAAGTGTGTCACCGTCCGAGTCCGACTCATGAGAAGATCATAC 1057

Db 241 ACAGTACGTCGAGGAGAAAGTGTGTCACCGTCCGAGTCCGACTCATGAGAAGATCATAC 300

QY 1058 CTAATTTGGGAATTTGCCACAGACAGTTATGACATTTGGGTTTATTTTGAATCG 1117

Db 301 CTA-ATTTGGGAATTTGCCACAGACAGTTATGACATTTGGGTTTATTTTGAATCG 357

QY 1118 ACAGACTCTCCAAATGCTCTGTCAGTGTGCATGTCAGTCCAGTCCAGTCCAGGAGGAG 1177

Db 358 ACAGACTCTCCAAATGCTCTGTCAGTGTGCATGTCAGTCCAGTCCAGTCCAGGAGGAG 405

QY 1178 GAGGAGGAGAAATGCTCACTTGTGAAGAAAAAGCAAAAAAGCAACCAACAGCCTCTG 1237

Db 406 GAGGAGGAGAAATGCTCACTTGTGAAGAAAAAGCAAAAAAGCAACCAACAGCCTCTG 465

QY 1238 CTGGATGAGATTGTACTGTGTACCGCGGGAGCTGTCA CGAGAAGTATATGACGACGC 1297

Db 466 CTGGATGAGATTGTACTGTGTACCGCGGGAGCTGTCA CGAGAAGTATATGACGACGC 525

QY 1298 CACAGTATCCAGGAGGGAGTCTATCTCTCAAGTTTGAATAATTCCTACTCTCTGTGG 1357

Db 526 CACAGTATCCCA-GGAGGGAGTCTATCTCTCAAG-TTGATAATTCCTACTCTCTGTGG 583

QY 1358 AGGTCCCAAGTCGCTCTACTACAGAGTCTATTACTAGATAGAGCTGTGTTCCAAAGTC 1417

Db 584 AGGTCCCAAGTCGCTCTACTACAGAGTCTATTACTAGATAGAGCTGTGTTCCAGGCTC 643

QY 1418 CGAGTCCAGGG-TTGAGCACAAACATGACGTTTAATTCCTTT 1459

Db 644 GCGAGTCTAGGGTTTGAGCACAAACATGACGTTTAATTCCTTT 686

## RESULT 13

CK627880

LOCUS

DEFINITION

CK627880

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: [graeme@helix.nih.gov](mailto:graeme@helix.nih.gov)

Plate: 04 row: a column: 04

Seq primer: M13RPL reverse primer (ABI).

Location/Qualifiers

1..603

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="ip04a04"

/tissue\_type="Whole eye"

/dev\_stage="Adult"

/lab\_host="EMDH10B"

/clone\_lib="Mouse Whole eye, unamplified: io/ip"

/note="Organ: Eye; Vector: pSport1; Approximately 1mg

total RNA was extracted from 200 adult mouse whole eyes. A

directionally cloned cDNA library in the pSPORT1 vector

(Invitrogen) was constructed at Bioserve Biotechnology

(Laurel MD) essentially following the protocols of the

SuperScript Plasmid System full details of which are

contained in the manufacturer's instruction manual

(<http://www.lifetechn.com/>). First strand synthesis was

carried out using a Not I primer-adaptor

15'-pGACTAGTCTAGATCGGAGCGGCCCTT15-3'. cDNA was

cloned in Not I/Sal I sites. EST analysis was performed on

the unamplified library at the NIH Intramural Sequencing

Center (NISC)."

## ORIGIN

Query Match 40.1%; Score 585.4; DB 7; Length 603;

Best Local Similarity 98.2%; Pred. No. 4e-139;

Matches 592; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 338 GTTGGTCCCAAGATAGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 397

Db 1 GTTGGTCCCAAGATAGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

QY 398 CGAAGGACGCGTGAAG 457



Db 61 CGAAGGCGGCTGAGAGGAGAGACGAGACGGCTGCAAAAGGAAGAGAGCGGAAG 120  
Qy 458 CGAGAGGAGGAACCGGCTGAGACGGAGGAGGAGAGAGCGCGGATAGAGGAGAG 517  
Db 121 CGAGAGGAGGAACCGGCTGAGACGGAGGAGGAGAGAGCGCGGATAGAGGAGAG 180  
Qy 518 AGGCTTCGGCTGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 577  
Db 181 AGGCTTCGGCTGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
Qy 578 GTGCAATTCAGCAGTATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 637  
Db 241 GTGCAATTCAGCAGTATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
Qy 638 CTCATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 697  
Db 301 CTCATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
Qy 698 CAACCTGCGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 757  
Db 361 CAGCTTGGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
Qy 758 TCATTCGCTGCATCATCAAGGTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 817  
Db 421 TCATTCGCTGCATCATCAAGGTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
Qy 818 GGAAGGCGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 877  
Db 481 GGAAGGCGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
Qy 878 GAAGCCTTGGAAATGGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 937  
Db 541 GAAGCCTTGGAAATGGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
Qy 938 ACA 940  
Db 601 ACA 603

RESULT 14  
AL546667  
LOCUS  
DEFINITION  
AL546667 Homo sapiens PLACENTA COT 25-NORMALIZED EST 25-MAR-2004  
clone CSODI029YJ06 5-PRIME, mRNA sequence.

ACCESSION  
AL546667  
VERSION  
EST.  
KEYWORDS  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:31268500.  
Contact: Genoscope

Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
962.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CSODI029YJ06&p1sc=962.r.

FEATURES  
Source

1. .1087  
/organism="Homo sapiens"  
/mol\_type="mRNA"

RESULT 15  
CN310226

/db\_xref="taxon:9606"  
/clone="CSODI029YJ06"  
/issue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 39.5%; Score 576.6; DB 1; Length 1087;  
Best Local Similarity 83.2%; Pred. No. 8.9e-137;  
Matches 651; Conservative 12; Mismatches 118; Indels 1; Gaps 1;  
Qy 97 AGTAAAGATGGCAAGCCCTTTTCATCAACTATGAAGAAAACCTGAAGTTCTGGCACT 156  
Db 288 AGAAGAGATGGCAAGCATTTTCATCAACTATGAAGAAAACCTGAAGTTCTGGCACT 347  
Qy 157 GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCCTGAGGTTGGATTCTT 216  
Db 348 GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTTCCTGAGTTGGATTCTT 407  
Qy 217 TGATGTTGGGGAATGATAGAGGAGAGAAATGGGAGACTCTGGGAAAACATGTCCAGGA 276  
Db 408 TGATGTTGGGGAATGATAGAGGAGAGAAATGGGAGACTCTGGGAAAACATGTCCAGGA 467  
Qy 277 GCATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAGTTGTTCCTCTCTCGGCATA 336  
Db 468 GCATGCCATGGTAGAGTTTGTGAAGCTTCTAAATAGTTGTTCCTCTCTCGGCATA 527  
Qy 337 TGTTCGCTCCCAAGAAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 396  
Db 528 TGTTCGCTCCCAAGAAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 587  
Qy 397 GCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 456  
Db 588 GCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 647  
Qy 457 GCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 516  
Db 648 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 707  
Qy 517 GAGGCTTCGGCTGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 576  
Db 708 AAGGCTTCGGCTGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 767  
Qy 577 CGTCAATTCAGCAGTATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636  
Db 768 CGTCAATTCAGCAGTATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 827  
Qy 637 TCTCATCGCCAGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 696  
Db 828 TCTCATCGCCAGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 887  
Qy 697 CCAACCTGCACAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 756  
Db 888 CCAGCTTCGACAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 947  
Qy 757 ATCATTCGCTGCATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCACTTAA 816  
Db 948 TTCTTCCTGCCTACATCATCAAGGTGAACACAGCTGGAGCAAGTATACACTGTCACTTAA 1007  
Qy 817 TGGACAGGCGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 876  
Db 1088 TGGACAGGCGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1066  
Qy 877 AG 878  
Db 1067 AV 1068

LOCUS CN310226 792 bp mRNA linear EST 16-MAY-2004  
DEFINITION 17000424186898 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN310226  
VERSION CN310226.1 GI:47326640  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 792)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R.,  
Lebkowski, J. and Stanton, L. W.

TITLE Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)  
COMMENT Contact: Brandenberger R

Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 792 Std Error: 0.00.

FEATURES  
source  
1..792  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/note="oligo dT primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

## ORIGIN

Query Match 39.5%; Score 576.4; DB 7; Length 792;  
Best Local Similarity 85.9%; Pred. No. 9e-137;  
Matches 663; Conservative 0; Mismatches 106; Indels 3; Gaps 2;  
QY 95 CTAGTAAAGATGGGAAAGCCTTCATCCAACTTATGAGAAAGAACTGAAGTTCGTGGCA 154  
DB 21 CAAGAAAGATGGCAAGACATTTTCATCCAACTTATGAGAAAGAAATTTGAAGCTTGTGGCA 80  
QY 155 CTGCATAAGCAAGTCTTTTGGGCCATATACCCAGACACGTCCTCGAGGTTGGATTTC 214  
DB 81 CTGCATAGCAAGTCTTTATGGGCCATATATCCAGACACTTGTCTCGAGGTTGGATTTC 140  
QY 215 TTTGATGTTGGGGAATGATAGGAGGAGGAATGGCGAGCTCTGGGAAACATGTCCAAG 274  
DB 141 TTTGATGTTGGGGAATGACAGGAGGAGGAATGGCGAGCTCTGGGAAACATGTCTAAA 200  
QY 275 GAGGATGCCATGGTAGAGTTGTGAAGCTTCTAAATAAGTGTGCTCTCTCTCTCGGCA 334  
DB 201 GAGGATGCCATGGTAGAGTTGTGAAGCTTCTAAATAAGTGTGCTCTCTCTCTCTCAACA 260  
QY 335 TATGTTGCTGCCACAGATAGAGAAGGAAGAG--AGAAAGAAAGAGCGGAGG 392  
DB 261 TATGTTGCTGCCACAAATAGAGAAGGAAGCAAGAAAGAAAGGAAGGGAAG 320  
QY 393 AGGAGCGAAGGCGCTGAAGGGAAGAACGAGAGCGGCTGCAGAAAGGAAGGAAGC 452  
DB 321 AGGAGCGAAGGCGCTGAAGGGAAGAACGCTCTGCAGAAAGGAAGGAAGAAC 380  
QY 453 GGAAGCGAGAGGAGGAGACCGGCTGAGCGGAGGAGGAAGAGGCGCGGATAGAG 512  
DB 381 GTAGGAGAGGAAGAGGAGGCTTCGACGGGAGGAAGGAAGGAGGAGCGGATAGAAG 440  
QY 513 AAGAGAGGCTTCGGCTGGGAACAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCGCAGA 572  
DB 441 AAGAAAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATATGGCAGCTTTAAACTCCAGA 500

QY 573 CTGCGCTGCAATTCAGCAGTATGCGAGCCAGCAGTATCCAGGGAACTACGAAACACAGC 632  
DB 501 CTGCGCTGCAATTCAGCAGTATGCGAGCCAGCAGTATCCAGGGAACTACGAAACACAGC 560  
QY 633 AGATTCTCATCGCCAGCTGCGAGGAGCAGCAGTATCAGCAGTATAAAACACAGGCGAGAGC 692  
DB 561 AAATTCTCATCGCCAGTTGCGAGGAGCAACACTATCAGCAGTACATGCGAGCAGTTGTATC 620  
QY 693 AAACCCAACTGCACAAACACAGGCGAGCATTTACAGAAACAGCAAGAGTAGTGGCTG 752  
DB 621 AAGTCCAGCTTGACAGCAACAGGCGAGCATTTACAGAAACACAGGAAAGTAGTAGTGGCTG 680  
QY 753 GGGCATCATGCTGCTCATCATCAAAAGGTGAACACAGCTGGAGCAAGTATACACTGTGAG 812  
DB 681 GGTCTTCTTGCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCTAG 740  
QY 813 TTAATGGACA-GGCCAAAACCCACACTGAAAATTCGAAAAAGTCTCTTGAGC 863  
DB 741 TTAATGGACAGGGGCCAAAACACACACTGACAGCTCGAAAAAGAACTGGAAC 792

Search completed: August 23, 2005, 00:03:58  
Job time : 5124 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 23, 2005, 00:08:19 ; Search time 6200 Seconds  
(without alignments)  
3618.510 Million cell updates/sec

Title: US-09-762-594-7  
Perfect score: 2394  
Sequence: 1 RRRPKVELFTVVTRVKVL.....DNSYSLWRKSVYRVYTR 463

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFTO.spool/US09762594/runat\_19082005\_171348\_18319/app\_query.fasta\_1.647  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09762594 @CNC 1 1 3731 @runat\_19082005\_171348\_18319 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2204	92.1	1724	10 AF022770	AF022770 Mus muscu
2	2200	91.9	3405	10 BC060602	BC060602 Mus muscu
3	2166.5	90.5	3461	10 AY336075	AY336075 Rattus no
4	2155.5	90.0	1927	10 BC083877	BC083877 Rattus no

5	2046	85.5	3598	9	AB043587	AB043587 Homo sapi
6	2043	85.3	2140	9	AK025520	AK025520 Homo sapi
7	2043	85.3	3358	9	BC060792	BC060792 Homo sapi
8	2021	84.4	1481	6	BD261678	BD261678 Neurotran
9	1983	82.8	3572	9	AY150218	AY150218 Homo sapi
10	1931	80.7	3492	5	AJ720620	AJ720620 Gallus ga
11	1699.5	71.0	1598	4	AY644721	AY644721 Sus scrof
12	1676	70.0	3037	9	BC034563	BC034563 Homo sapi
13	1505	62.9	961	6	AX677275	AX677275 Sequence
14	1345	56.2	859	6	CQ714948	CQ714948 Sequence
15	1176	49.1	2714	6	BD242865	BD242865 Secreted
16	1151	48.1	789	3	CQ731146	CQ731146 Sequence
17	1086.5	45.4	1897	3	AK116435	AK116435 Clona int
18	851	35.5	1495	6	CQ606435	CQ606435 Sequence
19	851	35.5	1779	3	AY051848	AY051848 Drosophil
20	752.5	31.4	3817	2	CQ606434	CQ606434 Sequence
21	752.5	31.4	14679	2	AC018262	AC018262 Drosophil
22	752.5	31.4	180263	3	AC010671	AC010671 Drosophil
23	752.5	31.4	207432	3	AB003513	AB003513 Drosophil
24	608	25.4	18140	10	AF501319	AF501319 Mus muscu
25	558	23.3	178951	5	BX323819	BX323819 Zebrafish
26	534	22.3	193192	10	AC121292	AC121292 Mus muscu
27	514.5	21.5	1759	5	AJ720156	AJ720156 Gallus ga
28	509	21.3	882	6	AX400067	AX400067 Sequence
29	506	21.1	860	6	CQ720008	CQ720008 Sequence
30	506	21.1	978	9	AY598329	AY598329 Homo sapi
31	506	21.1	3547	6	AX833708	AX833708 Sequence
32	506	21.1	3547	9	AK095650	AK095650 Homo sapi
33	506	21.1	230768	2	AC140763	AC140763 Rattus no
34	506	21.1	240225	2	AC126290	AC126290 Rattus no
35	492	20.6	593	10	BC050121	BC050121 Mus muscu
36	476	19.9	99251	9	AL592045	AL592045 Human DNA
37	476	19.9	140409	2	AC044825	AC044825 Homo sapi
38	476	19.9	232180	2	AC021883	AC021883 Homo sapi
39	472	19.7	992	5	BC054676	BC054676 Danio rer
40	455.5	19.0	88723	5	BX510322	BX510322 Zebrafish
41	435.5	18.2	150641	3	CEY41E3	CEY41E3 Caenorhabdi
42	410.5	17.1	3656	6	CQ606431	CQ606431 Sequence
43	392.5	16.4	432	6	AR521842	AR521842 Sequence
44	392.5	16.4	569	6	AR506297	AR506297 Sequence
45	384	16.0	278	6	AX524297	AX524297 Sequence

ALIGNMENTS

RESULT 1	AF022770	1724 bp	mRNA	linear	ROD 13-SEP-2002
LOCUS	Mus musculus peripheral benzodiazepine receptor associated protein (pap7) mRNA, complete cds.				
DEFINITION	AF022770				
ACCESSION	AF022770.3	GI:22831366			
VERSION					
KEYWORDS					
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE					
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1. (bases 1 to 1724)				
TITLE	Li, H., Degenhardt, B., Tobin, D., Yao, Z. X., Taeken, K. and Papadopoulos, V.				
JOURNAL	Identification, localization, and function in steroidogenesis of PAP7: a peripheral-type benzodiazepine receptor- and PKA (Rialpha)-associated protein				
MEDLINE	Mol. Endocrinol. 15 (12), 2211-2228 (2001)				
PUBMED	21588728				
REFERENCE	2 (bases 936 to 1490)				
AUTHORS	Li, H. and Papadopoulos, V.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-SEP-1997) Department of Cell Biology, Georgetown University Medical Center, 3900 Reservoir Rd. NW, Washington, DC 20007, USA				
REFERENCE	3 (bases 1 to 1724)				





Db 951 TCATTGCTGCATCATCAAGGTGAACACAGTTGGAGCAAGTGATACATGTCAGTTAAT 1010  
Qy 271 GlyGluAlaLysThrHisThrGluAenSerGluLysValLeuGluProGluAlaGlu 290  
Db 1011 GGACAGGCCAAAACCCACACTGAAATTCGAAAAAGTCTTGAGCGAGAGCTGCAGAA 1070  
Qy 291 GluAlaLeuGluAenGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrp 310  
Db 1071 GAAGCCTTGAAATATCGAACAAAGACTCTCTCCAGTGATGACCTCCATCCATGTGG 1130  
Qy 311 ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIle 330  
Db 1131 ACBAGACCACAAATCAAGACTTTAAAGAGAGATTCCGAGGATCCAGATTCTGTGATT 1190  
Qy 331 ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr 350  
Db 1191 ACAGTACGTCGAGGAGAGTCGTCACCGTCCGAGTCCGACTCATGAGGAGGATCATAC 1250  
Qy 351 LeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrp 370  
Db 1251 CTATTTTGGGAATTTGCCACAGACAGATTATGACATTTGGGTTTTATTTTGAATGG 1310  
Qy 371 ThrAspSerProAsnAlaAlaValSerValHisValSerGluSerAspGluGluGlu 390  
Db 1311 ACAGACTCTCCAAATGCTGCTGTCAGTGTGCATGTGCTGAGTCCAGTCCAGGAGGAG 1370  
Qy 391 GluGluGluGluAenValThrCysGluGluLysAlaLysLysAsnAlaLysProLeu 410  
Db 1371 GAGGAGGAAGAAATGTCACCTGTGTGAAGAAAGCAAAAGAACGCCAACAGCCTCTG 1430  
Qy 411 LeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySer 430  
Db 1431 CTGGATGAGATTGTACTGTGTACCGCGCGGACTGTCCAGGAGAGATATATGAGGAGC 1490  
Qy 431 HisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrp 450  
Db 1491 CACCAGTATCCAGGAGGGAGTCTATCTCTCCAGTTGTATTAATTCCTACTCTGTGG 1550  
Qy 451 ArgSerLysSerValTyrTyrArgValTyrTyrThrArg 463  
Db 1551 AGGTCCAGTCCGCTACTACAGAGTCTATTATATACAGA 1589

RESULT 3  
AY336075 3461 bp mRNA linear ROD 27-JUL-2003  
LOCUS Rattus norvegicus DMT1-associated protein mRNA, complete cds.  
DEFINITION AY336075  
ACCESSION AY336075  
VERSION AY336075.1 GI:33114044  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 3461)  
AUTHORS Chen,Y., Rodriguez-Paris,J.M., Ma,Y., Yeh,M., Yeh,K.-Y. and  
Glass,J.  
TITLE Intestinal Iron Absorption: the Interaction of DMT1 with DAP, an  
Iron Responsive Protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3461)  
AUTHORS Chen,Y., Rodriguez-Paris,J.M., Ma,Y., Yeh,M., Yeh,K.-Y. and  
Glass,J.  
TITLE Direct Submission  
JOURNAL Submitted (03-JUL-2003) Feist-Weiller Cancer Center, LSU Health  
Sciences Center, 1501 Kings Highway, Shreveport, LA 71105, USA  
FEATURES Location/Qualifiers  
1..3461  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/country="Yugoslavia: Belgrade"  
59..1639  
CDS

/note="golgi resident protein GCP60; peripheral benzodiazepine receptor-associated protein; golgi phosphoprotein 1; PBR-associated protein; PKA (Rialpha)-associated protein; DAP"  
/codon\_start=1  
/product="DMT1-associated protein"  
/protein\_id="AAP94639.1"  
/db\_xref="GI:33114045"  
/translation="MAQLNVQLSEVLDLTLSPDSEERPGAEGLQTPSPSPPRD  
GIGSTAGQEPGEGAAEAGAEARMEQHMGFLLELYGLALRFYKIDQKARHPT  
YBEKLFVALHKOVLGPNPDTSPVGFDFVLGNDRRREWALGNMSKEDAMVEPVK  
LKNKCPILLSAVVASHRIEKEERKRKAEREROREERERLOKEERKRREEDR  
LRESEERRRTEERLRLLEOQKOQIQAALNSQTAQFQQAARQVPGNEQOQLIRQ  
LOEQHYQOQOOLYQVLAQOQAALQOQVAVGASLPASTKQNTAGASDPLPVNQ  
AKHTENPEKVLPEAEALBENPKDSLPIAAPSMMTRPQIKDFKEKIRQADSVI  
TVRRGEVVTVPTEHEGSLFWBFATDSDYDIGFVYFEMTDPNAVSVHVSSESD  
EBEEBEENTVEEKAKNVNKLFLDEIVPVVRRDCHEEVYAGSHQVPGRVYLLKPDN  
SYSLWRSKSVYRVYVYTR"

ORIGIN

Alignment Scores:  
Pred. No.: 1.27e-153 Length: 3461  
Score: 2166.50 Matches: 420  
Percent Similarity: 97.47% Conservative: 3  
Best Local Similarity: 96.77% Mismatches: 10  
Query Match: 90.50% Indels: 1  
DB: 10 Gaps: 1  
US-09-762-594-7 (1-463) x AY336075 (1-3461)  
Qy 31 ValLysAspGlyLysAlaPheHisProThrTyrGluLysLeuLysPheValAlaLeu 50  
Db 335 ATANAAGATGGCAACCCCTTCCTCAACTTACGAAAGAAAACTGAAGTTTGTGGCAGTG 394  
Qy 51 HisLysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePhe 70  
Db 395 CATAGCAAGTCTCTTTGGGCCCATATATACCCAGACACTTCCCTCGAGGTGGATTCTTT 454  
Qy 71 AspValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGlu 90  
Db 455 GATGTGTTGGGAATGATAGGAGGAGAGAAATGGCAGCTCTGGAAACATGTCCAAAGAG 514  
Qy 91 AspAlaMetValGluPheValLysLeuAsnLysCysProLeuLeuSerAlaTyr 110  
Db 515 GACGCCATGGTAGAGTTGTGAAGCTTCTGAATAAGTGTCTCTCTCAGCATAT 574  
Qy 111 ValAlaSerHisArgIleGluLysGluGluGluLysArgLysAlaGluGluGlu 130  
Db 575 GTTGGTCCCACAGATAG 634  
Qy 131 ArgArgGlnArgGluGluGluArgGluArgLeuGlnLysGluGluLysArgLys 150  
Db 635 CGAAGCGACGCTGAAGAGAAACGAGAGCGCTGCAGAAAGAGAGAGAGAGAGAGAGAG 694  
Qy 151 ArgGluGluGluAspArgLeuArgGluGluGluArgArgArgIleGluGluGlu 170  
Db 695 CGAGAGGAGGAGGACCGACTCAGACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 754  
Qy 171 ArgLeuArgLeuGluGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAla 190  
Db 755 AGGCTTCGGCTGGAACAGACAAAGCAGCAGATATGGCAGCTTTAACTCGCAGAGTGC 814  
Qy 191 ValGlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIle 210  
Db 815 GTGCAATTCACAGTATGCGGCCCCAGCAGTAGTATCCAGGAACTACGAGCAGCAGATT 874  
Qy 211 LeuIleArgGlnLeuGlnGlnIleHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThr 230  
Db 875 CTCATCCCGCAGTCGAGGAGCAGCAGTATCAGCAGTATACCAACAGTTATACCAAGTC 934  
Qy 231 GlnProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAla 250  
Db 935 CAGCTTGCACACACAGCAGCATTTGCAGAGAGCAGCAAGAGTAGTGTGTGGCTGGGCT 994

251 SerLeuProLalaSerSerSerValAsnThrAlaGlyAlaSerAspThrLeuSerValAsn 270  
 |||||  
 995 TCATTGCTGCGTCAACAAAGGTGAACACAGCTGGAGCAAGTGATCCCTGCGCAGTGAAC 1054  
 |||||  
 271 GlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGlu 290  
 |||||  
 1055 GGACAGCCCAAGACCCACATCGAATCCCGAAAAGTCTCTGAGCCAGAAAGCTGGGAA 1114  
 |||||  
 291 GluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaLaProSerMetTir 310  
 |||||  
 1115 GAGGCTTGGAATATGACCAAGACTCTCTCCAGTGATTCAGCTCTCTTCATGTGG 1174  
 |||||  
 311 ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIle 330  
 |||||  
 1175 ACAAGACCAACAGATCAAGAGCTTAAAGAGAGATTCGGCAGGATCGAGATTCTGTGATT 1234  
 |||||  
 331 ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr 350  
 |||||  
 1235 ACAGTCGGCGGAGAGAGTAGTCACCGTTTCGAGTCCCACTCATGAGGAAGATCATAC 1294  
 |||||  
 351 LeuPheTirPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTir 370  
 |||||  
 1295 CTCTTTTGGGAATTTGCCACAGACAGTTATGACATTGGGTTTGGGGTTTATTTGATGG 1354  
 |||||  
 371 ThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSer---AspGluGlu 389  
 |||||  
 1355 ACAGACTCTCCAAACGCTGTCGTGTCAGTGTGCATGTCAGTCCAGTGACGACGAGAG 1414  
 |||||  
 390 GluGluGluGluGluAsnValThrCysGluGluLysAlaLysIleAsnAlaAsnLysPro 409  
 |||||  
 1415 GAGGAGAGAGAGAAATGTCACTTGTGAAGAGAAAGCAAAAAGATGTCAACAAAGCCT 1474  
 |||||  
 410 LeuLeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGly 429  
 |||||  
 1475 CTCTCGATGAGATTGTACTGTGTACCGCGGAGTGTCTATGAGAGAGTGTATGAGGC 1534  
 |||||  
 430 SerHisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeu 449  
 |||||  
 1535 AGCCACCAAGTACCCAGGAGGGGGTCTATCTCTCAAGTTGATTAACCTTACTCTCTG 1594  
 |||||  
 450 TtrArgSerLysSerValTyrTyrArgValTyrTyrArg 463  
 |||||  
 1595 TGGAGGTCCCAAGTCAGTCTACTACAGAGTCTATTACTAGTA 1636  
 |||||

RESULT 4  
 BC083877  
 LOCUS  
 DEFINITION Rattus norvegicus DMT1-associated protein, mRNA ROD 12-OCT-2004  
 MGC:95164 IMAGE:7130229, complete cds.  
 ACCESSION BC083877  
 VERSION  
 KEYWORDS  
 SOURCE MGC.  
 ORGANISM Rattus norvegicus (Norway rat)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 1927)  
 Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Krausner,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
 Carninci,P., Prange,C., Raha,S.S., Loguercio,N.A., Peters,G.J.,  
 Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Boiffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,  
 Schnerch,A., Schein,J.B., Jones,S.J. and Marra,M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 1927)  
 Director MGC Project.  
 Direct Submission  
 Submitted (01-Oct-2004) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs-femail.nih.gov](mailto:cgabbs-femail.nih.gov)  
 Tissue Procurement: Howard Jacobs  
 cDNA Library Preparation: Express Genomics  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 184 Row: h Column: 2  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 33469058.

## FEATURES

Source

1..1927  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="MGC:95164 IMAGE:7130229"  
 /issue\_type="Kidney, rat (Brown Norway)"  
 /clone\_lib="NIH MGC 236"  
 /lab\_host="DH10B"  
 /note="Vector: pExpress1"  
 1..1927  
 /gene="Acdb3"  
 /db\_xref="LocusID:289312"  
 /db\_xref="RGD:727851"  
 20..1600  
 /gene="Acdb3"  
 /codon\_start=1  
 /product="DMT1-associated protein"  
 /protein\_id="AAH83877.1"  
 /db\_xref="GI:54035576"  
 /db\_xref="LocusID:289312"  
 /db\_xref="RGD:727851"  
 /translation="MAAQLNVQLVSLDGLTLSPDSEERPGAEGAPLOTPPSPPGD  
 GLGSTAGAQOREPGEAAAGAAEABERMEHQHGFGLLELYGLALRYKIKDGAHPPT  
 YBEKLKVALHKQVLGTPNPDTSPGVGFVIGNDRRREWAALGNMSKEDAWVEFVK  
 LNKCCPLLSAYVASHRIKTEBEERKAEERREERREERREERREERREERREERRE  
 LRREERRRIEERLEQQQIMAAINSQTAVQFOQYAAQYQYEQQILIRQ  
 LQEQHYQYMQQLYQVLAQQAALQKQEVVVAGASPAKAKVNTAGADPLPVNGQ  
 AKTHTENPKVLEPAEAELNGPKDSLPTVAAPSMWTRPQIKDFKEKTRQADSVI  
 TYRGEVTVYRVPTHEEGSYLWFEEATDSYDYGFGVYFETWSPNAAVSHVHSESSD  
 EEEEEEVTCBEKAKNVNKLDELIVPVYRDRDCHEEVYAGSHQYPPGRGVLKLFEDN  
 SYSLWRSKSVYRVYVYTR"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,2e-153 Length: 1927  
 Score: 2155.50 Matches: 418  
 Percent Similarity: 97.00% Conservative: 3  
 Best Local Similarity: 96.31% Mismatches: 12  
 Query Match: 90.04% Indels: 1



DB:	10	Gaps:	1
US-09-762-594-7 (1-463) x BC083877 (1-1927)			
QY	31	ValLysAspGlyLeuAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeu	50
Db	296	ATAAAGATGGCAAGCCTTCATCCAACTTACGAGAGAAAACGAAAGTTTGGCAGCTG	355
QY	51	HisLysGlnValLeuLeuGlyProTyrAenProAepThrSerProGluValGlyPhePhe	70
Db	356	CATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACTTCCTCCCTGAGGTGGATTCTTT	415
QY	71	AspValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGlu	90
Db	416	GATGTGTTGGGGAATGATAGGAGGAGAGATGGCAGCTCTGGGAAACATGTCCAAGAG	475
QY	91	AspAlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyr	110
Db	476	GAGCCATGTTAGAGTTTGTGAAGCTTCTGAATAGTGTCTCTCTCTCAGCATAT	535
QY	111	ValAlaSerHisArgIleGluLysGluGluGluLysLeuLysArgLysAlaGluGlu	130
Db	536	GTTGCTCCACAGAAATGAGAAGGAAGAAAGAAAGAAAGAAAGCGGAGGAGGAG	595
QY	131	ArgArgGlnArgGluGluGluGluArgGluGluGluGluGluGluGluGluGlu	150
Db	596	CGAAGCAGCGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCGGAAG	655
QY	151	ArgGluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGlu	170
Db	656	CGAGGGAGAGACCGACTCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	715
QY	171	ArgLeuArgGluGluGlnGlnLysGlnGlnIleMetAlaLeuAsnSerGlnThrAla	190
Db	716	AGCTTCGGCTGGAACAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCGCAGCTGC	775
QY	191	ValGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIle	210
Db	776	GTGCAATTCAGCAGTATCGGCCAGCAGTATCCAGGGAATCTACGAGCAGCAGATT	835
QY	211	LeuIleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThr	230
Db	836	CTCATCCCGCAGTCGAGGAGCAGCAGCTATCAGCAGTACATGCACAGTTATACCAATC	895
QY	231	GlnProAlaGlnGlnGlnAlaLeuGlnLysGlnGlnValValMetAlaGlyAla	250
Db	896	CAGCTTGCCAGCAACAGCAGCAGTTCGAGAAGCAGCAAGAGTAGTGTGGTGGGCT	955
QY	251	SerLeuProAlaSerSerLysValAlaThrAlaGlyAlaSerAspThrLeuSerValAsn	270
Db	956	TCAATTCCTGCGTCGCAAGAGTGAAACACAGCTGGAGCAGGTGATCCCTGCGAGTGAAC	1015
QY	271	GlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGlu	290
Db	1016	GGACAGCCCAAGACCCACACTGAGATCCCGAGAAGTCTCTGAGCCAGAGCTCGGAA	1075
QY	291	GluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyr	310
Db	1076	GAGGCCCTGGAAATGGACCCCAAGACTCTCTCCAGTGAATTCAGCTCCTTCATGTGG	1135
QY	311	ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIle	330
Db	1136	ACAAGACCACATCAAGAGACTTTAAAGAGAGATTCGCGAGATGTCAGATTCTGTGATT	1195
QY	331	ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr	350
Db	1196	ACAGTCCGCGAGGAGAGTAGTACCGTTCGAGTCCCACTCATGAGAGAGATCATAC	1255
QY	351	LeuPheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyr	370
Db	1256	CTCTTTTGGGAATTTGCCACAGACAGTATGATGATGTTGGGTTTATTTTGAATGG	1315
QY	371	ThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSer---AspGluGlu	389

```

3'UTR      1643..3580
           /gene="gcp60"

ORIGIN
Alignment Scores:
Pred. No.:      1,616-144      Length:      3598
Score:          2046.00
Percent Similarity: 94.91%
Best Local Similarity: 90.28%
Query Match:      85.46%
DB:               9
Gaps:             0

US-09-762-594-7 (1-463) x AB043587 (1-3598)

QY      32  LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
DB      344 AAAGATGGCAAGCAATTCATCCAACTTATGAGAGAAAATTTGAAGCTTGTGGCAGTCGAT 403

QY      52  LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
DB      404 AACCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTGGATTCTTTGAT 463

QY      72  ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91
DB      464 GTGTTGGGGAATCACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGCTTAAAGAGGAT 523

QY      92  AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111
DB      524 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGGTGTGGCATCTCTTTTCAACATATGTT 583

QY      112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGluArg 131
DB      584 GCGTCCCAAAATAGAGGAAGAGCAAGCAAGAAAAAGGAAGGAGGAGGAGGAGCGA 643

QY      132 ArgGlnArgGluGluGluGluArgGluGluGluGluGluGluGluGluGluGluGlu 151
DB      644 AGCGGCGGTGAAGAGGAAGAGAGACGCTCTCAAAAGAGGAGGAGGAGGAGGAGGAGGAG 703

QY      152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluGluGlu 171
DB      704 GAAGAGAGGAAGAGGCTTCGACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 763

QY      172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
DB      764 CTTGCGTTGAGCAGCAAAAGCAGCAGATAATGCGCAGCTTTAACTCCCAAGCTGCGCGT 823

QY      192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211
DB      824 CAGTCCAGCAGTATGACGCCCAACAGTATCCAGGGAACCTACGAACAGCAGCAAAATCTC 883

QY      212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
DB      884 ATCCGCGAGTTGCAGGAGCAACACTATCAGCAGTATCAGCAGCTGTTGATCAAGTCCAG 943

QY      232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGlnValMetAlaGlyAlaSer 251
DB      944 CTTGCACAGCAACAGCAGCATTACAGAAACACAGCAAGTAGTAGTGGCTGGCTCTCC 1003

QY      252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
DB      1004 TTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAGTTATGGA 1063

QY      272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
DB      1064 CAGGCCAAAACACACTGACAGCTCCGAAAAGAACTGGAACCAAGCAAGCTGCAAGAA 1123

QY      292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
DB      1124 GCCCTGGAGAAATGGACCAAAAGAAATCTCTCCAGTAATAGCAGCTCCATCCATGGACA 1183

QY      312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
DB      1184 CGACCTTCAGATCAAGACTTCAAGAGAGAGATTTCAGCAGGATGCGATTCGCTGATTACA 1243

```

```

QY      332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
DB      1244 GTGGCCGAGGAGAGAGTGGTCACTGTTGAGTACCCACCCATGAAGAGGATCATATCTC 1303

QY      352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThr 371
DB      1304 TTTTGGGAATTTGCCACAGACAATATGACATTTGGGTTTGGGGTGTATTTTGAATGGACA 1363

QY      372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
DB      1364 GACTCTCCAACACTGCTGTGACGGTGCATGTGAGTGAGTCCAGCGCATGACGAGGAG 1423

QY      392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAlaLysAlaAsnLysProLeuLeu 411
DB      1424 GAAGAGAAAACATCGTGTGAGAGAGAAAGCCAAAAGAAATGCCAACAGCTTTGCTG 1483

QY      412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
DB      1484 GATGAGATTGTGCTGTGTACCGCGGACTGTGTATGAGGAGGTGTATGCTGCGCAGCAT 1543

QY      432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTyrArg 451
DB      1544 CAATATCCAGGAGAGAGGTATCTCTCAAGTTTGCAAACTCTTACTCTTTGTGGCGG 1603

QY      452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
DB      1604 TCAAAATCAGTCTACTACAGAGTCTATTATACATAGA 1639

RESULT 6
AK025520      2140 bp      mRNA      linear      PRI 13-SEP-2003
LOCUS      Homo sapiens cDNA: FLJ21867 fis, clone HEP02419.
ACCESSION      AK025520
VERSION      AK025520.1 GI:10438060
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
            NEDO human cDNA sequencing project
            Unpublished.
TITLE      2 (bases 1 to 2140)
JOURNAL      Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
            Direct Submission
            Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT      NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan: cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Department of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES      Location/Qualifiers
            1..2140
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="HEP02419"
            /cell_line="HepG2"
            /cell_type="hepatoma"
            /note="cloning vector pME18SFL3"
            49..1635
            CDS

```



human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 REFERENCE  
 2 (bases 1 to 3358)  
 AUTHORS  
 Strausberg, R.  
 TITLE  
 Direct Submission  
 Submitted (03-NOV-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 133 Row: 1 Column: 24  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 15826851.

#### FEATURES

Location/Qualifiers  
 1..3358  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:71593 IMAGE:4813993"  
 /tissue\_type="Brain, hippocampus"  
 /clone\_lib="NIH MGC\_95"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"  
 1..3358  
 /gene="ACBD3"  
 /note="synonyms: GCP60, PAP7, GOCAP1, GOLPH1"  
 /db\_xref="LocusID:64746"  
 /db\_xref="MIM:606809"  
 44..1630  
 /gene="ACBD3"  
 /codon\_start=1  
 /product="ACBD3 protein"  
 /protein\_id="AAH60792.1"  
 /db\_xref="GI:38511512"  
 /db\_xref="LocusID:64746"  
 /db\_xref="MIM:606809"  
 /translation="MAVLNARLEVSVDGLTSLDPDPERPGAGAPLPPPLPPSP  
 PSCGPGAGSQPEGAAGAAAEARRLEQRWGFLEELYGLALRPFKEKDGKAF  
 HPTBEKLALVHLQVLMGPNPDTPCVGFDFVLGNDRREREAALGNMSKDAMVE  
 FVLLNRCCLFSTVYASHKIEKEOEKKRKEBERRRERERLQKEERKREE  
 EERLREERERERLEERLQKQQLQVQALNSQVQFQYAAQVPGNYEQQIL  
 IRLQEQHYQYMQYLQVQLAQQLQKQVAVVAGSLFTSSKVNATVPSNMVSV  
 NQAKHTDSSKELEPEAAEALENGPKESLPVIAAPSMWTRPQIKDKIKQDAD  
 SVITVGRGVVTVPTVEGSLFWFEPATNDIDFGYFVFWTDSPTNAVSVHVS  
 SDDDEEENIGCEKAKKNANRPLLDLIVPVYRRDCHEVYAGSHQYVGRGVLLKF  
 DNSYSLWRSKSVYRYVYTR"  
 353..550  
 /gene="ACBD3"  
 /note="ACBP; Region: Acyl CoA binding protein"  
 /db\_xref="CDD:pfam00887"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 2,496-144 Length: 3358  
 Score: 2043.00 Matches: 389  
 Percent Similarity: 94.91% Conservative: 21  
 Best Local Similarity: 90.05% Mismatches: 22

Query Match: 85.34% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-762-594-7 (1-463) x BC060792 (1-3358)  
 QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
 332 AAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCAGTCAT 391  
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71  
 392 AAGCAAGTTCCTTATGGCCCATATAATCCAGACACTTGTCTCGAGGTGGATTCCTTGTAT 451  
 QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyValAsnMetSerLysGluAsp 91  
 452 GTGTGGGGAATGACAGGAGGAGAGATGGGAGCCCTGGGAAACATGTCTTAAGAGGAT 511  
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111  
 512 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGGTGTGCCATCTCTTTTCAACATATGTT 571  
 QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGluArg 131  
 572 GCGTCCCAAAATAGAGGAAGCAAGCAAGAAAAAAGGAAGGAGGAGGAGGAGCGA 631  
 QY 132 ArgGlnArgGluGluGluGluArgGluArgLeuGlnLysGluGluLysArgLysArg 151  
 632 AGCGCGCTGAAGAGGAGAGAGAGAACTCTGCAAGAGGAGGAGGAGGAGGAGGAGGAG 691  
 QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgIleGluGluGluArg 171  
 692 GAAGAAGAGGAAAGGCTTCGACGGGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 751  
 QY 172 LeuArgLeuGluGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191  
 752 CTTGGTGGAGCAGCAAAAGCAGCAGATATATGGCAGCTTTTAACTCCAGAGTCGCGTG 811  
 QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211  
 812 CAGTTCACAGCATATGACGCCCAACAGTATCCAGGAACTACGACGAGCAAAATTC 871  
 QY 212 IleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231  
 872 ATCCGCCAGTTTCAGGAGCAACACTATCAGCAGTATCAGCAGCAGTGTGTATCAAGTCCAG 931  
 QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251  
 932 CTTGCACAGCAACAGCAGCAGCATTACAGAAACCAACAGAAAGTAGTAGTGGCTCTTCC 991  
 QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271  
 992 TTGCTCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGATGATTAATGGA 1051  
 QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291  
 1052 CAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAAACCAAGAGCTGCAGAGAA 1111  
 QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311  
 1112 GCCTCGAGATGGCAAAAGAAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACA 1171  
 QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAspSerValIleThr 331  
 1172 CCACTCTAGATCAAGACTTCAAGAGAGAGATTCAGCAGGATGCAGATTCGGGTGATTACA 1231  
 QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351  
 1232 GTGGGCCGAGGAGAGTGGTCACTGTTCCAGTACCCACCACCAAGAGAGATCATATCTC 1291  
 QY 352 PheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThr 371  
 1292 TTTTGGGAATTTGCCACACACAAATATGACATGGGTTGGGTTGTTATTTGAATGGACA 1351



Accession	Protein	Start	End	Score	E-value	Ident	Mismatch	Gap	Frame
Db		1214				GATGAGATTGTCGCTGTGTATCCGACGGGACTGT			
						CATGAGAGGTGTATGTCGCAGCAT			
Qy		432				GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrrpArg			
Db		1274				CAATATCCAGGAGAGAGAGGTCTATCTCTCAAGTTTGACAACTCCCTACTCTTTGTGGCGG			

Qy 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463  
|||  
Db 1334 TCAAATCAGTCTCTACAGAGTCTATATACTAGA 1363

RESULT 9			
AY150218			
LOCUS	AY150218	3572 bp	mRNA
DEFINITION	Homo sapiens peripheral benzodiazepine receptor associated protein mRNA, complete cds.		linear PRI 21-APR-2003

## ORIGIN

Alignment Scores:		
Pred. No.:	8,91e-140	Length:
Cred.	1983.00	Matches:
Score	385	Conservative:
Percent Similarity:	94.00%	Mismatches:
Best Local Similarity:	88.91%	Indels:
Query Match:	82.83%	Gaps:
DB:	9	

US-09-762-594÷7 (1-463) x AY150218 (1-3572)

**Qy**    32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
         |||||  
**Db**    315 AAAGATGGCAAAGCATTTCTCAACTATTGAAGAAAAATTGAAGCTTGTCGCACGTGAT 378  
         |||||



QY 411 uapGluIleValProValTyArgArgAspCysHisGluGluValTyAlaGlySerHi 431  
Db 1454 GGATGAGATTGCTCTGTGTACCGACGGAGCTGTCTATGAGGAGGTGTATGCTGGCAGCCA 1513  
QY 431 sGIntyrProGlyArgGlyValTyArgGlyLeuLeuLysPheAspAsnSerTyrSerLeuTrpAr 451  
Db 1514 TCAATATCCAGGAGGAGGTCTATCTCTCAAGTTTGACAACTCTCTCTCTTTGTGGCG 1573  
QY 451 gSerLysSerValTyTyArgValTyTyArg 463  
Db 1574 GTCAAAATCAGTCTACTACAGAGTCTTATTATATACAGA 1610  
RESULT 10  
AJ720620 3492 bp mRNA linear VRT 30-SEP-2004  
LOCUS Gallus gallus mRNA for hypothetical protein, clone 21m6.  
DEFINITION  
ACCESSION AJ720620  
VERSION  
AJ720620.1 GI:53133899  
KEYWORDS  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE  
1 Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezubov, Y., Zaim, J.,  
Fiedler, P., Kutter, S., Blagodatski, A., Kostovska, D., Koter, M.,  
Plachy, J., Carninci, P., Hayashizaki, Y. and Buerstedde, J.M.  
Full-length cDNAs from bursal lymphocytes to facilitate gene  
function analysis  
Unpublished  
2 (bases 1 to 3492)  
Caldwell, R.B.  
Direct Submission  
Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,  
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.  
1, D-85764 Neuherberg, GERMANY  
FEATURES  
Location/Qualifiers  
source  
1..3492  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="CB"  
/db\_xref="taxon:9031"  
/clone="21m6"  
/cell\_type="lymphocyte"  
/tissue\_type="bursa"  
/clone\_lib="rikeni"  
/dev\_stage="2 weeks old"  
1..3492  
/locus\_tag="RCJMB04\_21m6"  
<1..50  
/locus\_tag="RCJMB04\_21m6"  
51..1610  
/locus\_tag="RCJMB04\_21m6"  
/note="ORF1"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAG32279.1"  
/db\_xref="GI:53133900"  
/translation="MAAALSARLEVSVDGLTLPNAEAPPCEARPGHGDSSAAGNRRA  
GGCPSOAEGEEEAALSPERWGFALBELVGLALRPFKEKDGAFHPTYEKRLV  
ALHKVQLVDPNDTPCEVDFDVLGNDRRKEWAALGNMSKOEAMTFVKLNRCHL  
FSTYVSHKIEKEBEKEERREERREERLERLQKEERQOEERLEERLEER  
RRIEERLEMEQKQKIWAALNSQTAMQFQYAAQYQYGNYPQOQILIRLOEYHQ  
YMQYLYQVLAQQALQKQEAQVAAAGTPLTTASKVNPVQGMPSINGQASHTD  
NPEKELDPEALEENGPKDSVPVTAAPSMARPOIKDFEKIRQDADSVITVGRGE  
VTVVTPHTEGSLYFWFATSDYDIGCVFEWTDSPNTAVSHVSSSDDEDEE  
NASSBEKAKKANQQLDEILVLYRDRCHEEVYAGSHQYVPGRGVILLKPDNSYLSLRS  
KTVYRVYVYTR"  
1611..3492  
/locus\_tag="RCJMB04\_21m6"

Alignment Scores:  
Pred. No.: 7,18e-136 Length: 3492  
Score: 1931.00 Matches: 368  
Percent Similarity: 90.55% Conservative: 25  
Best Local Similarity: 84.79% Mismatches: 39  
Query Match: 80.66% Indels: 2  
DB: 5 Gaps: 1  
US-09-762-594-7 (1-463) x AJ720620 (1-3492)  
QY 32 LysAspGlyLeuAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
Db 306 AAAGATGGCAAGCCTTTCATCAACATATGAAGAAAACACTCAGACTTGTGGCACTGCAC 365  
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71  
Db 366 AAGCAGGTTCTGCTGGACCTTTACAACCCCTGATCTTGCCTCAAGCTGGTCTTTGTGAT 425  
QY 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
Db 426 GTGCTGGGGAATGATAGAGGAAGGATGGGCTGGCCCTTGGAAACATGTGGAAGCAAGAA 485  
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111  
Db 486 GCTATGACAGAAATTTGTTAAGCTCTTAATAGTGTGCTGCACCTTGTTCACAGTATGTT 545  
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgIleGluAlaGluGluArg 131  
Db 546 ACTTCTTCACAGATAGAAAAGAAACAAGAAAAGAAAGAGAGAGGAGGAGGAAACGA 605  
QY 132 ArgGlnArgGluGluGluArgGluGluGluLysGluGluLysArgLysArg 151  
Db 606 AGGCGCCGTGAAGAAGAGAGGCTTGAGCGGTTTACAAAAGAAAGAAAACGCGAGAA 665  
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgIleGluGluArg 171  
Db 666 GAAGAAGAAGAGAGGCTGGAAGAGAGGAGAGAGAGAGAGAGGAGGAGGAGCGG 725  
QY 172 LeuArgLeuGluGlnLysGlnIleMetAlaLeuAsnSerGlnThrAlaVal 191  
Db 726 CTTCCGATGGAAACAGCAAAAGCAACAGATCATGGCAGCACTCAACTCCACAGATGCCATG 785  
QY 192 GlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnLeu 211  
Db 786 CAGTTCACAGATGCTGCTCAGCAGATATCCCGCACTATGAACAGCAGCAGATATCCTC 845  
QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231  
Db 846 ATTCGACAGCTCCAGAGCAGCAGCTATCAACAGTACATGCGCAGCAGTGTGATCAAGTCCAG 905  
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251  
Db 906 CTTGCACAGCAACAGCAGCCTTGCAGAAAACAGCAGGAGGCGAGTGTGTGGCAGCAGCGG 965  
QY 252 LeuPro-----AlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerVal 269  
Db 966 ACACCTCTGACTACTGATCCCAAGGTGAATGTACTCTGCCAAGGGGATATGCCGTCTATT 1025  
QY 270 AsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAla 289  
Db 1026 ANTGGCAAGCCAGTGCACACACAGACAACCCCTGAAAAGAGAGCTGGATCCAGAGGCTTGG 1085  
QY 290 GluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMet 309  
Db 1086 GAAGAAGCATTTGGAAATGGACCAAAAGATTCTGTTCAGTGATAGCTGCTCCATCGATG 1145  
QY 310 TrpThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerVal 329  
Db 1146 TGGGCACGAGCCCGAGATAAAAGACTTCAAAGAGAAAATCCCGCAGATGCAGACTCTGTG 1205  
QY 330 IleThrValArgGlyGluValValThrValArgValProThrHisGluGluGlySer 349  
Db 1206 ATCAGATGGGCCGAGGGGAAGTGGTTACAGTTAGAGTACCAACTCATGAAGAGGGGTCT 1265







294	Qy	LeuGluProGluAlaAlaGluGluAlaLeuGluAenGlyProLysAapSerLeuProVal	303
541	Db	CTGGAACCAGAGGCTGCAGAGAAGCCCTCGAGAAATGGACCAAAAGAATCTCTTCCAGTA	600
304	Qy	IleAlaAlaProSerMetTrpThrArgProGlnIleLysAapPheLysGluLysIleArg	323
601	Db	ATAGCAGCTCCATCCATGTGCACAGACCTCAGATCAAGACTTCAAGAGAGAAGATTACG	660
324	Qy	GlnAspAlaAapSerValIleThrValArgArgGlyGluValValThrValArgValPro	343
661	Db	CAGGATGCAGATTCCGTGATTACAGTGGCCGAGGAGNAGTGTGCATCTTTCAGTACCC	720
344	Qy	ThrHisGluGluGlySerTyrLeuPheTrpGluPheAlaThrAapSerTyrAspIleGly	363
721	Db	ACCCATGAAGAAGGATCATATCTCTTTTGGAAATTGGCCACAGACAATTATGACATTTGGG	780
364	Qy	PheGlyValTyrPheGluTrpThrAapSerProAsnAlaAlaValSerValHisValSer	383
781	Db	TTTGGGGTGTAATTTTGAATGCACAGACTCTCAAACACGTGTGTGCAGCGTGCATGTCAGT	840
384	Qy	GluSerSerAapGluGluGluGluGluGluAenValThrCysGluGluLysAlaLys	403
841	Db	GAGTCCACGGATGACGACGAGAGAGAAGAAACATCGGTTGTGAAGAGAGAAGCCAAA	900
404	Qy	LysAsnAlaAenLysProLeuLeuAapGluIleValProValTyrArgArgAapCysHis	423
901	Db	AAGAATGCCAACAGCCCTTTGCTGGATGAGATTGTGCTGTGTACCGACGGGACTGTTCAT	960
424	Qy	GluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTyrLeuLeuLysPhe	443
961	Db	GAGGAGGTGTATGCTGGCAGCATCAATATCCAGGGAGAGAGTCTATCTCCTCAAGTTT	1020
444	Qy	AspAsnSerTyrSerLeuTrpArgSerLysSerValTyrTyrArgValTyrTyrThrArg	463
1021	Db	GACAACTCCTACTCTTTGTGGCGGTCAAAATCACTCTACTACAGAGCTATTATACATAGA	1080

RESULT 13					
AX677275					
LOCUS	AX677275	Sequence 53 from Patent WO02086122.	961 bp	DNA	linear
DEFINITION	AX677275				
ACCESSION	AX677275.1				
VERSION	AX677275.1	GI:29334680			

## ORIGIN

Alignment Scores:		
Pred. No.:	1.89e-104	361
Score:	1505.00	289
Percent Similarity:	94.98%	Conservative: 14
Best Local Similarity:	90.60%	Mismatches: 16
Query Match:	62.87%	Indels: 0
DB:	6	Gaps: 0

US-09-762-594-7 (1-463) x AX677275 (1-961)

**Qy**      145   GluGluGluLysArgGluGluGluAspArgLeuArgArgGluGluGluGluArg    164  
             |||||     |||||     |||||     |||||     |||||     |||||     |||||     |||||  
**Db**      2      GAGGAAGAGAAAACGTATGGAGAGAACAAGAGGAAAGCTTCACCGGAGGAGAGAGGAAAGG    61

QY	165	ArgArgIleGluGluGluAargLeuArgLeuGlnGlnLysGlnLysGlnLysMetAlaAla	184
Db	62	AGA CGG AT TAG AAG AAG AAC G C T T C G G T T G G AGC AGC AAA AAG CAG CAG ATAATATGGCAGCT	121
QY	185	LeuAsnSerGlnThrAlaValAlinPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsn	204
Db	122	T T A A A C T C C C A G A C T G C G T G C A G T T C C A G C A G T A T C C A G C C A A C A G T A T C C A G G A A C	181
QY	205	TyrGluGlnGlnLysLeuLeuLeuArgGlnLeuGlnGlnHisTyrGlnGlnTyrLys	224
Db	182	T A C G A A C A G C A G C A A A T T C T C A T C C G C C A G T T G C A G G A G C A A C A C A T A T C A G C A G T A C A T G	241
QY	225	HisGlnAlaGluGlnThrGlnProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnGlu	244
Db	242	C A G C A G T T G T A T C A A G T C C A G C T T T G C A C A C A C A G G C A G C A T T A C A G A A A C A A C A G G A A	301
QY	245	ValValMetAlaGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSer	264
Db	302	G T A G T A G G C T G G G T C T T C T C G C C T A C A T C A A A G T G A A T G C A A C T G C A A C T G T A C C A A G T	361
QY	265	AspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeu	284
Db	362	A A T A T G A T G T C A G T T A A T G A C A G C G C C A A A C A C A C A C A C A C T G A C A C T C C G A A A A A A A A C T G	421
QY	295	GluProGluAlaAlaGluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIle	304
Db	422	G A A C C A A A G C T G C A G A A G A A G C C C T G G A A T T G G A C C A A A A G A A T C T C T T C C A G T A A T A	481
QY	305	AlaAlaProSerMetTrpThrArgProGlnIleLysAspPhelyAspGluLysIleArgGln	324
Db	482	G C A C T C A T C C A T G T G A C A C A G A C C T C A G A T C A A G A C T T C A A G A G A A G A A T T C A G C A G	541
QY	325	AspAlaAspSerValIleThrValArgArgGlyGluValValThrValArgValProThr	344
Db	542	G A T C A G A T T C C G T G A T T A C A G T G G G C G A G G A A G T G T C A C T G T T C G A G T A C C A C C	601
QY	345	HisGluGluGlySerTyrLeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPhe	364
Db	602	C A T G A A G A A G A T C A T A T C T C T T T G G A A T T T G C C A C A G A C A A T T A T G A C A T T G G G T T T	661
QY	365	GlyValTyrPheGluTrpThrAspSerProAsnAlaAlaValSerValHisValSerGlu	384
Db	662	G G G T G T A T T T G A A T G A C A G A C T C T C C A A C A C T G C T G C A G C G T G C A T G T C A G T A G	721
QY	385	SerSerAspGluGluGluGluGluGluUasnValThrCysGluGluLysAlaLysLys	404
Db	722	T C C A G C A T A C A G C A G A G A G A A A A A C A T C G G T G T G A G A G A A A G C C A A A A A G	781
QY	405	AsnAlaAsnLysProLeuLeuAspGluIleValProValTyrArgArgAspCysHisGlu	424
Db	782	A A T C C C A A C A A G C C T T T G C T G G A T G A G A T T G C C C T G T P A C C A G C G G A C T G C A T A G	841
QY	425	GluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAsp	444
Db	842	G A G T G T A T G T G C A G C C A T C A A T A T C C A G G A G A G A G T C T A T C T C C T C A A G T T T G A C	901
QY	445	AsnSerTyrSerLeuTrpArgSerLysSerValTyrTyrArgValTyrTyrThrArg	463
Db	902	A A C T C T C T C T T T G T G G C G T C A A A A T C A G T C T A C A G A G T C T A T T A T A C T A G A	958
<hr/>			
RESULT 14			
LOCUS CQ714948			
DEFINITION Sequence 882 from Patent WO02068579.			
ACCESSION CQ714948			
VERSION CQ714948.1 GI:42275805			
KEYWORDS Homo sapiens (human)			
SOURCE Homo sapiens			
ORGANISM linear PAT 03-FBB-2004			
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1			

AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof  
JOURNAL Patent: WO 02068579-A 882 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES Location/Qualifiers  
Source 1..859  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores: 1.88e-92 Length: 859  
Pred. No.: 1345.00 Matches: 257  
Score: 94.39% Conservative: 12  
Percent Similarity: 90.18% Mismatches: 16  
Best Local Similarity: 56.18% Indels: 0  
Query Match: 6 Gaps: 0  
DB: 6  
US-09-762-594-7 (1-463) x CQ714948 (1-859)  
QY 179 GlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyrAlaAla 198  
DB 2 CAGCAGATATGGCAGCTTAACTCCAGAGTGGCGTGCAGTTCAGCAGTATGCAGCC 61  
QY 199 GlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeuIleAargGlnLeuGln 218  
DB 62 CAACAGTATCCAGGGAATACGAACAGCAGCAAAATCTCATCCGCCAGTTGCAGGAGCAA 121  
QY 219 HisTyrGlnGlnTyrIleHisGlnAlaGluGlnThrGlnProAlaGlnGlnAlaAla 238  
DB 122 CACTATCAGCAGTACATGCAGCAGTGTATCAAGTCCAGCTTGCACAGCAACAGGAGCA 181  
QY 239 LeuGlnLysGlnGlnGlnValMetAlaGlyAlaSerLeuProAlaSerSerLysVal 258  
DB 182 TTACAGAAACACAGAGAGTAGTAGTGGCTGGTCTTCCTTGCCTCATCATCAAAAGTG 241  
QY 259 AsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlnAlaLysThrHisThrGlu 278  
DB 242 AATGCAACTGTACCAAGTAAATATGATGTAGTGTAGTGGTCTTCCTTGCCTCATCATCAAAAGTG 301  
QY 279 AsnSerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyProLys 298  
DB 302 AGCTCCGAAAAAGAACTGGAACAGAGCTGCAGAGAGAGCCCTGGAGAAATGCACAAA 361  
QY 299 AspSerLeuProValIleAlaProSerMetTyrThrArgProGlnIleLysAspPhe 318  
DB 362 GAATCTCTCCAGTATAGCAGCTCCATCCATGTCGACAGCAGCTCAGATCAAGACTTC 421  
QY 319 LysGluLysIleAargGlnAspAlaAspSerValIleThrValAargGlyGluValVal 338  
DB 422 AAAGAGAAGATTCAGCAGGATGCAGATTCGGTGATTACAGTGGCGCCAGGAGAGTGC 481  
QY 339 ThrValArgValProThrHisGluGluGlySerTyrLeuPheThrGluPheAlaThrAsp 358  
DB 482 ACTGTTTCGAGTACCCACCCATGAAGAGGATCATATCTCTTTGGGAATTTGCCACAGAC 541  
QY 359 SerTyrAspIleGlyPheGlyValTyrPheGluTyrThrAspSerProAsnAlaVal 378  
DB 542 AATTATGACATTTGGGTGTATTTTGAATGACAGACTCTCCAAACACTGCTGTC 601  
QY 379 SerValHisValSerGluSerAspGluGluGluGluGluGluGluGluGluGluGlu 398  
DB 602 AGCGTCCATGTCAGTCCAGCGATGACGACGAGGAGGAGAGAGAGAGAGAGAGAGAG 661  
QY 399 GluGluLysAlaLysLysAlaAsnLysProLeuLeuAspGluIleValProValTyr 418  
DB 662 GAAGAGAAAGCCAAAAAGAAATGCAACAGCCCTTTGCTGGATGAGATTGCGCTGTGTAC 721  
QY 419 AtgArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyVal 438

DB 722 CGAGGGAGTGTATGATGAGAGGTGTATGCTGGCAGCCATCAATATCCAGGAGAGAGTC 781  
QY 439 TyrLeuLeuLysPheAspAsnSerTyrSerLeuTyrArgSerLysSerValTyrTyrArg 458  
DB 782 TATCTCTCAAGTTTGACAACCTCTACTCTTTGTGGCGGTCAAAATCAGTCTACTACAGA 841  
QY 459 ValTyrTyrTyrArg 463  
DB 842 GTCTATTATACTAGA 856  
RESULT 15  
BD242865  
LOCUS 2714 bp DNA linear PAT 17-JUL-2003  
DEFINITION Secreted proteins and polynucleotides encoding them.  
ACCESSION BD242865  
VERSION BD242865.1 GI:33052635  
KEYWORDS JP 2002536973-A/16.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2714)  
AUTHORS Valenzuela, D., Yuan, O., Hoffman, H., Hall, J. and Rapiejko, P.  
TITLE Secreted proteins and polynucleotides encoding them  
JOURNAL Patent: JP 2002536973-A 16 05-NOV-2002;  
COMMENT ALPHAGEN INC  
OS Homo sapiens (human)  
PN JP 2002536973-A/16  
PD 05-NOV-2002  
PF 18-FEB-2000 JP 2000599860  
PR 19-FEB-1999 US 60/120680, 23-APR-1999 US 09/298733 PR  
17-AUG-1999 US 60/149639, 23-SEP-1999 US 60/155686 PR  
01-OCT-1999 US 60/157247, 29-NOV-1999 US 60/167823 PR  
29-NOV-1999 US 60/167822, 15-FEB-2000 US 60/182711 PI DARIO  
VALENZUELA, OLIVE YUAN, HEIDI HOFFMAN, JEFF HALL, PETER PI RAPIEJKO  
PC C12N15/09, A61K38/00, A61P3/10, A61P5/14, A61P11/00, A61P11/06, PC  
A61P19/02  
PC A61P21/04, A61P25/14, A61P27/02, A61P29/00, A61P31/04, A61P31/10,  
PC A61P31/12,  
PC A61P31/18, A61P31/20, A61P31/22, A61P37/00, A61P37/06, C07K14/435,  
PC C12N5/10,  
PC C12P19/34, C12R1/91, C12N15/00, C12N5/00, A61K37/02  
CC Secreted proteins and polynucleotides encoding them FH Key  
Location/Qualifiers  
FT source 1..2714  
FT Location/Qualifiers  
1..2714  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores: 4.15e-79 Length: 2714  
Pred. No.: 1176.00 Matches: 224  
Score: 93.23% Conservative: 11  
Percent Similarity: 88.83% Mismatches: 17  
Best Local Similarity: 49.12% Indels: 0  
Query Match: 6 Gaps: 0  
DB: 6  
US-09-762-594-7 (1-463) x BD242865 (1-2714)  
QY 212 IleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231  
DB 2 ATCCGCCAGTTCAGGAGCAACACTATCAGCAGTACATGCAGCAGTGTGTATCAAGTCCAG 61  
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGluValValMetAlaGlyAlaSer 251  
DB 62 CTTGCACAGCAACAGCGCAGCATTCAGAAACAGGAAAGTAGTAGTGGCTGGTTTCC 121  
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271

```
Db      122 TTGCTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTGAGTTAATGGA 181
Qy      272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db      182 CAGGCCAAAACACACACTGACAGCTCCGAAAAGAACTGGACCAAGCTGCAGAAGAA 241
Qy      292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
Db      242 GCCTCGAGAAATGGACCAAGAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACA 301
Qy      312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db      302 CGACCTCAGATCAAGACTTCAAGAGAAAGATTTCAGCAGGATGCGATTCCGGATTACA 361
Qy      332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
Db      362 GTGGGCGAGGAGAGTGGTCACTGTTTCGAGTACCCACCACCATGAAGAAGGATCATATCTC 421
Qy      352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
Db      422 TTTTGGGAATTTGCCACAGACAATTATGACATTGGGTTTGGGGTGTATTTTGAATGGACA 481
Qy      372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
Db      482 GACTCTCCAAACACTGCTGTACGGTGCATGTGAGTGAGTCCAGCGGATGACGACGAGGAG 541
Qy      392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
Db      542 GAAGAGAAAACATCGGTTGTGAAGAGAAAGCCAAAAGAAATGCCAACAGCCTTTGCTG 601
Qy      412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
Db      602 GATGAGATTGTGCTGTGTACCGACGGGACTGTCTGAGGAGGTGTATGCTGGCAGCCAT 661
Qy      432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
Db      662 CAATATCCAGGGAGAGAGGCTATCTCCTCAAGTTTGACAACTCCTACTCTTTTGTGGCGG 721
Qy      452 SerLysSerValTyrTyrArgValTyrTyrArg 463
Db      722 TCAAAATCAGTCTACTACAGAGTCTATTATTACTAGA 757
```

Search completed: August 23, 2005, 04:17:41  
Job time : 6220 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 23, 2005, 00:04:03 ; Search time 750 seconds

(without alignments)  
3654.454 Million cell updates/sec

Title: US-09-762-594-7

Perfect score: 2394

Sequence: 1 RPRRPKVELFTVTVTRKVVVL.....DNSYSLMRSKSVYRVYTR 463

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q/cn2\_1/USPTO spool/US09762594/runat\_19082005\_171348\_18291/app.query.fasta\_1.647  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09762594 @CGN 1.1 470 @runat\_19082005\_171348\_18291 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004as:.\*  
13: Geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2394	100.0	1459	3 AAZ57038	Aaz57038 PBR-assoc
2	2046	85.5	3560	13 ADS34290	Ads34290 POSH prot
3	2046	85.5	3598	13 ADS34286	Ads34286 POSH prot
4	2046	85.5	3598	13 ADS34287	Ads34287 POSH prot
5	2043	85.3	2140	13 ADS34285	Ads34285 POSH prot

6	2042	85.3	3399	4 AAK52051	Aak52051 Human pol
7	2033	84.9	1995	13 ADS34291	Ads34291 POSH prot
8	2021	84.4	1481	3 AAZ87207	Aaz87207 Human NTA
9	2002.5	83.6	2350	4 AAK53035	Aak53035 Human pol
10	1983	82.8	3572	13 ADS34289	Ads34289 POSH prot
11	1676	70.0	3049	13 ADS34288	Ads34288 POSH prot
12	1505	62.9	961	8 ACA56933	Acas6933 Human adi
13	1176	49.1	2714	3 AAA93116	Aaa93116 Human sec
14	851	35.5	1495	4 ABL24635	Ab124635 Drosophil
15	752.5	31.4	3817	4 ABL24634	Ab124634 Drosophil
16	558	23.3	420	3 AAC77363	Aac77363 Human ORF
17	509	21.3	882	6 ABQ93525	Abq93525 Human cdn
18	506	21.1	3547	11 ADM02147	Adm02147 Human cdn
19	442	18.5	291	3 AAC77466	Aac77466 Human ORF
20	410.5	17.1	3656	4 ABL24632	Ab124632 Drosophil
21	384	16.0	278	4 AAK53762	Aak53762 Murine tr
22	376	15.7	2450	8 ABZ73815	Abz73815 Secreterd
23	376	15.7	2450	10 ABZ67399	Abz67399 Human sec
24	352	14.7	934	6 ABR24346	Abk24346 DNA encod
25	309	12.9	210204	12 ADQ18927	Adq18927 Human sof
26	304.5	12.7	554	4 AAI17988	Aai17988 Probe #79
27	304.5	12.7	554	4 ABA62955	Abas62955 Human foe
28	304.5	12.7	554	4 AAI42978	Aai42978 Probe #11
29	304.5	12.7	554	4 AAK37154	Aak37154 Human bon
30	304.5	12.7	554	4 AAK11364	Aak11364 Human bra
31	304.5	12.7	554	4 ABS36831	Abs36831 Human liv
32	304.5	12.7	554	6 ABS11145	Abs11145 Human gen
33	302	12.6	306	4 AAI28147	Aai28147 Probe #18
34	302	12.6	306	4 ABA76496	Aba76496 Human foe
35	302	12.6	306	4 AAI57184	Aai57184 Probe #25
36	302	12.6	306	4 AAK51149	Aak51149 Human bon
37	302	12.6	306	4 AAK25143	Aak25143 Human bra
38	302	12.6	306	4 ABS50709	Abs50709 Human liv
39	302	12.6	306	6 ABS24667	Abs24667 Human gen
40	267.5	11.2	4184	4 ABL24630	Ab124630 Drosophil
41	267.5	11.2	4384	4 AAK77371	Aak77371 Human imm
42	241	10.1	873	4 AAH72906	Aah72906 Human cer
43	230.5	9.6	4959	5 AAH81756	Aah81756 Human dif
44	230.5	9.6	5810	4 AAK52571	Aak52571 Human pol
45	230.5	9.6	5810	4 AAK52572	Aak52572 Human pol

ALIGNMENTS

RESULT 1  
AAZ57038  
ID AAZ57038 standard; cdna; 1459 BP.  
XX  
AC AAZ57038;  
XX  
DT 19-MAY-2000 (first entry)  
XX  
DE PBR-associated protein (PAP)7 encoding cdna.

XX Peripheral-type benzodiazepine receptor; PBR; PBR- associated protein;  
KW PAP; cell proliferation; cancer; cell death; cytostatic; neuroprotective;  
KW immunomodulator; antiinfertility; cerebroprotective; atherosclerosis;  
KW Niemann-Pick C; tumour; Alzheimer's disease; developmental disorder;  
KW cholesterol; multiple sclerosis; stress; neurodegenerative disorder;  
KW immune disorder; stroke; PAP7; ss.  
XX  
XX Mus sp.

XX Key Location/Qualifiers  
FH CDS 8..1399  
FT /\*tag= a  
FT /product= "PAP7"

XX WO200009549-A2.

XX 24-FEB-2000.

XX 11-AUG-1999; 99WO-US018507.



XX 11-AUG-1998; 98US-0096048P.  
 PR (GEOU ) UNIV GEORGETOWN MEDICAL CENT.  
 XX  
 XX Papadopoulos V, Li H;  
 XX WPI: 2000-224278/19.  
 DR P-PSDB; AA67500.  
 DR  
 XX Novel peripheral-type benzodiazepine receptor associated proteins used  
 PT for the regulation of the peripheral-type benzodiazepine receptor.  
 PT  
 XX Claim 3; Page 68; 71pp; English.  
 XX  
 XX The invention provides isolated peripheral-type benzodiazepine receptor  
 CC (PBR) associated proteins (PAPs) and nucleic acids encoding the PAP  
 CC proteins. The PAP polynucleotides are a source of primers and probes for  
 CC detection, isolation and amplification. PAP ligands or substrates or  
 CC antibodies can be labeled and used to detect PAPs, in the diagnosis or  
 CC prognosis of disease associated with increased cell proliferation, such  
 CC as cancer, or reduced cell death. The diagnostic methods of the invention  
 CC can be predictive of diseases involving PBR including gallstones,  
 CC atherosclerosis, Niemann-Pick C, Sitosterolemia, Dystrophy, tumor  
 CC proliferation, Schnyder's corneal crystalline dystrophy, brain disorders  
 CC including Alzheimer's disease, cholesterol metabolism, Tellurium  
 CC toxicity, Smith-Lemli-Opitz syndrome, myelinization, developmental  
 CC abnormalities, demyelination, Charcot-Marie tooth disease, Pelizaeus-  
 CC Merzbacher disease, Multiple sclerosis, and SIA. The methods may also be  
 CC useful in prophylactic treatments, or in screening for compounds  
 CC effective in prophylactic treatment. The PAPs may be used to identify  
 CC inhibitors or activators which allows the identification of drugs or  
 CC agents which modulate PBR activity. Inhibitors of PAP may be used in the  
 CC treatment or amelioration of conditions such as stress and stroke,  
 CC cancer, neurodegenerative disorders, developmental disorders, infertility  
 CC and immune disorders. The present sequence represents a cDNA encoding a  
 CC PAP7 polypeptide  
 XX  
 SQ Sequence 1459 BP; 455 A; 298 C; 403 G; 303 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,62e-213 Length: 1459  
 Score: 2394.00 Matches: 463  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0  
 US-09-762-594-7 (1-463) x AA257038 (1-1459)  
 QY 1 ArgProArgArgProLysValGluLeuPheThrValValThrArgVallyValValLeu 20  
 DB 8 CGGCGCGCGCGACCTAAAGTTGAGTTGTTCTAGTACCGCGTGTGAAGGTAGTTTFA 67  
 QY 21 PheLeuAsnGlnLeuSerLeuCySLysLeuValLysAspGlyLysAlaPheHisProThr 40  
 DB 68 TTTTAAATCAACTTTCATTGTGCAAACTAGTAAAGATGGCAAGCCCTTTCATCAACT 127  
 QY 41 TyrGluGluLysLeuLysPheValAlaLeuHisLysGlnValLeuLeuGlyProTyrAsn 60  
 DB 128 TATGAAGAAAACCTGAAGTTCTGGGCACTGCATAAGCAAGTTCTTTTGGGCCCATATAAC 187  
 QY 61 ProAspThrSerProGluValGlyPhePheAspValLeuGlyAsnAspArgArgGlu 80  
 DB 188 CCAGACACGCTCCCTCAGGTTGGATTCTTTGATGTGTGGGAATGATAGGAGGAGAA 247  
 QY 81 TrpAlaAlaLeuGlyAsnMetSerLysGluAspAlaMetValGluPheValLysLeuLeu 100  
 DB 248 TGGGCAAGCTCTGGGAAACATGTCCAGAGAGATGCCATGTGTAGAGTTGTGAAGCTTCTA 307  
 QY 101 AsnLysCysCysProLeuLeuSerAlaTyrValAlaSerHisArgIleGluLysGluGlu 120  
 DB 308 AATAAGTGTGTCTCTCTCTCGGCATATGTTGTCCTCCACAGAAATAGAGAAGGAAGAA 367

QY 121 GluGluLysArgArgLysAlaGluGluArgGlnArgGlnGluGluGluGluGluGlu 140  
 DB 368 GAAGAGAAAAGAAAGAGCGGAGGAGGAGCGGAGCGGTGAAGAGGAAGACAGAG 427  
 QY 141 ArgLeuGlnLysGluGluLysArgLysArgGluGluGluGluGluGluGluGlu 160  
 DB 428 CGGCTGCAAAAGGAAGAGAGCGGAAAGCGAGAGGAGGAAGACCGGCTGAGACGGAG 487  
 QY 161 GluGluGluArgArgIleGluGluGluArgLeuArgLeuGluGluGlnGlnLysGlnGln 180  
 DB 488 GAGGAAGAGAGCGCGGATAGAGGAAGAGAGGCTTCGGCTGGAACAGCAAAAGCAGCAG 547  
 QY 181 IleMetAlaAlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyrAlaAlaGlnGln 200  
 DB 548 APTATGGCAGCTTTAAACTCGCAGACTGCCGTCGCAATTCACGACGATATGCAGCCAGCAG 607  
 QY 201 TyrProGlyAsnTyrGluGlnGlnGlnIleLeuIleArgGlnLeuGlnGlnGlnHisTyr 220  
 DB 608 TATCCAGGGAAGTACGAACAACAGCAGATCTTCATCCGCCAGCTGCAGGAGCAGCACTAT 667  
 QY 221 GlnGlnTyrLysHisGlnAlaGluGlnThrGlnProAlaGlnGlnGlnAlaAlaLeuGln 240  
 DB 668 CAGCAGTATAAACACACGAGCAGCAAAACCCAACTGCACAAACAAACAGCAGCATTTACAG 727  
 QY 241 LysGlnGlnGluValValMetAlaGlyAlaSerLeuProAlaSerSerLysValAlaThr 260  
 DB 728 AAACAGCAAGAAAGTAGTAGTGGCGCATCATTTGCCCTGCATCATCAAGGTGAACACA 787  
 QY 261 AlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsnSer 280  
 DB 788 GCTGGAGCAAGTGATACACTGTCAGTTAATGGACAGGCCAAACCCACACTGAAATTC 847  
 QY 281 GluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluGluGluGluGluGlu 300  
 DB 848 GAAAAGTCTTTGAGCCAGAGCTGCAGAAGAGCGCTTGGAAAATGGACAAAGACTCT 907  
 QY 301 LeuProValIleAlaAlaProSerMetTyrThrArgProGlnIleLysAspPheLysGlu 320  
 DB 908 CTTCCAGTGATTCAGCTCCATCCATGTGGACAACACCAACAATCAAGACTTTAAAGAG 967  
 QY 321 LysIleArgGlnAspAlaAspSerValIleThrValArgArgGlyGluValValThrVal 340  
 DB 968 AAGATTCCGACGATGCAGATTCTGTGATTACAGTACGTCGAGAGGAGAGTCTGTCCCGTC 1027  
 QY 341 ArgValProThrHisGluGluGlySerTyrLeuPheThrGluPheAlaThrAspSerTyr 360  
 DB 1028 CGAGTCCCGACTCATGAGGAAGGATCATACCTATTTTGGGAATTTGCCACAGACAGTTAT 1087  
 QY 361 AspIleGlyPheGlyValTyrPheGluTyrThrAspSerProAsnAlaAlaValSerVal 380  
 DB 1088 GACATTGGTTGGGGTTTATTTTGAATGGACAGACTCTCCAAATGTCTGTCTGAGTGTG 1147  
 QY 381 HisValSerGluSerSerAspGluGluGluGluGluGluGluGluGluGluGluGlu 400  
 DB 1148 CATGTGAGTGTCCAGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1207  
 QY 401 LysAlaLysLysAsnAlaAsnLysProLeuLeuAspGluIleValProValTyrArgArg 420  
 DB 1208 AAAGCAAAAAGAACCCCAAGGCTCTGTCTGGATGAGATTGTACCTGTGTACCGCGG 1267  
 QY 421 AspCysHisGluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTyrLeu 440  
 DB 1268 GACTGTACGAGGAAGTATATGACGAGCAGCCACAGTATCCAGGAGGAGGAGGAGTCTATCTC 1327  
 QY 441 LeuLysPheAspAsnSerTyrSerLeuTrpArgSerLysSerValTyrTyrArgValTyr 460  
 DB 1328 CTCAGTTTGATAAATTCCTACTCTCTGTGGAGGTCCAAAGTCCGTCTACTACAGAGTCTAT 1387  
 QY 461 TyrThrArg 463  
 DB 1388 TATACTAGA 1396

## RESULT 2

ADS34290  
ID ADS34290 standard; DNA; 3560 BP.

AC ADS34290;

XX DT 02-DEC-2004 (first entry)

XX DE POSH protein associated DNA #44.

XX ds; gene; cytostatic; neurotropic; neuroprotective; antiparkinsonian;  
KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI;  
KW Ubiquitin ligase; POSH-associated protein; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.

XX OS Homo sapiens.

XX PN WO2004078130-A2.

XX PD 16-SEP-2004.

XX PF 02-MAR-2004; 2004WO-US006308.

XX PR 03-MAR-2003; 2003US-0451437P.

XX PR 05-MAR-2003; 2003US-0452284P.

XX PR 19-MAR-2003; 2003US-0455760P.

XX PR 20-MAR-2003; 2003US-0456640P.

XX PR 03-APR-2003; 2003US-0460526P.

XX PR 04-APR-2003; 2003US-0460792P.

XX PR 21-APR-2003; 2003US-0464285P.

XX PR 05-MAY-2003; 2003US-0469462P.

XX PR 15-MAY-2003; 2003US-0471378P.

XX PR 20-MAY-2003; 2003US-0472327P.

XX PR 30-MAY-2003; 2003US-0474706P.

XX PR 03-JUN-2003; 2003US-0475825P.

XX PR 17-JUN-2003; 2003US-0479317P.

XX PR 19-JUN-2003; 2003US-0480215P.

XX PR 19-JUN-2003; 2003US-0480376P.

XX PR 28-AUG-2003; 2003US-0493860P.

XX PR 28-AUG-2003; 2003US-0498634P.

XX PR 16-SEP-2003; 2003US-0503931P.

XX PR 10-NOV-2003; 2003WO-US035712.

XX PR 05-FEB-2004; 2004WO-US003600.

XX PR 02-MAR-2004; 2004US-0549896P.

XX PA (PROT-) PROTEOLOGICS INC.

XX PI Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;

XX PI Greener T;

XX XX WPI; 2004-662346/64.

XX Isolated, purified or recombinant complex, useful for identifying an  
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
PT POSH-associated protein (POSH-AP).

XX PS Disclosure; SEQ ID NO 54; 374pp; English.

CC POSH-associated disease is viral infection, POSH-associated cancer or  
CC POSH-associated neurological disorder. The methods are useful for  
CC treating or preventing POSH-associated neurological disorder in a subject  
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
CC nucleic acid of the invention.

XX SQ Sequence 3560 BP; 1059 A; 661 C; 820 G; 1020 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,32e-180 Length: 3560  
Score: 2046.00 Matches: 390  
Percent Similarity: 94.91% Conservative: 20  
Best Local Similarity: 90.28% Mismatches: 22  
Query Match: 85.46% Indels: 0  
DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34290 (1-3560)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
DB 322 AAAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAATTGAAGCTTGTGGCAGTCAT 381  
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71  
DB 382 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTGGATTCTTTGAT 441  
QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
DB 442 GTCTTGGGAATGACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCTAAGAGGAT 501  
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111  
DB 502 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGTGTGTGCCATCTCTTTTCAACATATGTT 561  
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131  
DB 562 GCGTCCCAAAATAGAGGAAGCAAGCAAGAAAAAAGGAAGGAGGAGGAGGAGGAGGAG 621  
QY 132 ArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 151  
DB 622 AGCGCGGTGAAGAGGAAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAG 681  
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluGlu 171  
DB 682 GAAGAGAGAGAAAGGCTTCGACGGGAGGAAGAGAAAGAGAGAGAGAGAGAGAGAG 741  
QY 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191  
DB 742 CTTGGTTGGAGCAGCAAAAGCAGCAGAGATTAATGGCAGCTTTAAACTCCCAGACTGCGGTG 801  
QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211  
DB 802 CAGTTCCAGCAGTATGCAGCCCAACAGTAGTATCCAGGGAAGTACGAGCAGCAAAATTC 861  
QY 212 IleArgGlnGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231  
DB 862 ATCCGCCAGTTGCAGGAGCAACACTATATCAGCAGTATATCAGCAGGAGTTGTATCAAGTCCAG 921  
QY 232 ProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251  
DB 922 CTTGCACAGCAACAGCAGCAGCATTTACAGAAACACAGGAAGTAGTAGTGGCTGGCTTCC 981  
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271  
DB 982 TTTCCTTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGATGATTAATGA 1041  
QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLysGluGluProGluAlaAlaGluGlu 291  
DB 1042 CAGGCCCAAAACACACTGACAGCTCCGAAAGAAAGAACTGGAACCGGAAGCTGCAAGAA 1101  
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311

```
Db 1102 GCCTGAGAAATGGACCAAAAGAAATCTCTCCAGTAAATAGCAGCTCCATCCATGTGGACA 1161
Qy 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db 1162 CGACCTCAGATCAAGACTTCAAAGAGAGATTACAGCAGGATGCAGATTCGCTGATTACA 1221
Qy 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
Db 1222 GTGGCCGAGGAGGAATGGTCTCTGTCAGTACCCACCACCCATGAAGAGATCATATCTC 1281
Qy 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
Db 1282 TTTTGGGAATTTGCCACACACATATGACATTTGGGTGTTGGGTGTTATTTGAATGACA 1341
Qy 372 AspSerProAsnAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
Db 1342 GACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTCAGTCAGTCAGTCAGTCAGTCAG 1401
Qy 392 GluGluGluAsnValThrCysGluGluLysAlaLysAlaLysAlaLysAlaLysAlaLys 411
Db 1402 GAAGAAGAACAATCCGTTGTGAGAGAGAGCCAAAGAAAGTCCCAACAGCCCTTTGCTG 1461
Qy 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
Db 1462 GATGAGATTGCTGCTGTGTACCGACGGACTGTGTCAGGAGGTGTATGTCGACCCAT 1521
Qy 432 GlnTyrProGlyArgGlyValTyrLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
Db 1522 CAATATCCAGGAGAGAGAGTATCTCTCAAGTTTGACAACTCTCTACTCTTTGTGGCG 1581
Qy 452 SerLysSerValTyrArgValTyrArgValTyrThrArg 463
Db 1582 TCAAAATCAGTCTACTACAGAGTCTATTACTAGA 1617

RESULT 3
ID ADS34286
XX ADS34286 standard; DNA; 3598 BP.
AC ADS34286;
XX
XX
XX 02-DEC-2004 (first entry)
XX POSH protein associated DNA #40.
XX
KW de, gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
KW ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.
XX
OS Homo sapiens.
XX
XX WO2004078130-A2.
XX
XX 16-SEP-2004.
XX
XX 02-MAR-2004; 2004WO-US006308.
XX
XX 03-MAR-2003; 2003US-0451437P.
XX 05-MAR-2003; 2003US-0452284P.
XX 19-MAR-2003; 2003US-0455760P.
XX 20-MAR-2003; 2003US-0456640P.
XX 03-APR-2003; 2003US-0460526P.
XX 04-APR-2003; 2003US-0460792P.
XX 21-APR-2003; 2003US-0464285P.
XX 09-MAY-2003; 2003US-0469462P.
XX 15-MAY-2003; 2003US-0471378P.
XX 20-MAY-2003; 2003US-0472327P.
XX 30-MAY-2003; 2003US-0474706P.
XX 03-JUN-2003; 2003US-0475825P.
```

```
PR 17-JUN-2003; 2003US-0479317P.
PR 19-JUN-2003; 2003US-0480215P.
PR 19-JUN-2003; 2003US-0480376P.
PR 08-AUG-2003; 2003US-0493860P.
PR 28-AUG-2003; 2003US-0498634P.
PR 16-SEP-2003; 2003US-0503931P.
PR 10-NOV-2003; 2003WO-US035712.
PR 05-FEB-2004; 2004WO-US003600.
PR 02-MAR-2004; 2004US-0549896P.
XX
XX (PROT-) PROTEOLOGICS INC.
XX
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
XX Greener T;
XX
XX WPI; 2004-662346/64.
XX
XX Isolated, purified or recombinant complex, useful for identifying an
XX antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
XX POSH-associated protein (POSH-AP).
XX
XX Disclosure; SEQ ID NO 50; 374pp; English.
XX
XX The invention relates to an isolated, purified or recombinant complex (I)
XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
XX or HERPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are
XX useful for identifying an agent that modulates an activity of a POSH
XX polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
XX apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
XX of a protein through the secretory pathway, an agent that inhibits the
XX progression of a neurological disorder, an agent that modulates a POSH
XX function, an agent that modulates a HERPUD1 function. The methods can be
XX used for treating a viral infection, for inhibiting an activity of a POSH
XX -AP in a cell, for treating a POSH-associated disease in a subject. The
XX POSH-associated disease is viral infection, POSH-associated cancer or
XX POSH-associated neurological disorder. The methods are useful for
XX treating or preventing POSH-associated neurological disorder in a subject
XX e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
XX schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
XX nucleic acid of the invention.
XX
XX Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2,35e-180 Length: 3598
XX Score: 2046.00 Matches: 390
XX Percent Similarity: 94.91% Conservative: 20
XX Best Local Similarity: 90.28% Mismatches: 22
XX Query Match: 85.46% Indels: 0
XX DB: 13 Gaps: 0
XX
XX US-09-762-594-7 (1-463) x ADS34286 (1-3598)
Qy 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
Db 344 AAAGATGGCAAGCAATTCCTCAACTATGAAGAAAAAATTGAAGCTTGGCATTGCAT 403
Qy 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
Db 404 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCTGAGGTTCGATTTGAT 463
Qy 72 ValLeuGlyAsnAspArgArgGluTrpAlaLysLeuGlyAsnMetSerLysGluAsp 91
Db 464 GTGTTGGGGAATGACAGGAGGAGAGAGTGGGAGCCCTGGGAAACATGTCTAAAGAGGAT 523
Qy 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
Db 524 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGTGTTCATCTCTTTTCAACATATGTT 583
Qy 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
Db 584 GCGTCCCAAAATAGAGAGGAGAGCAAGAAAAAAGAGAGAGGAGGAGCGA 643
```



CC nucleic acid of the invention.

XX Sequence 3598 BP; 1071 A; 667 C; 826 G; 1034 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,35e-180 Length: 3598  
Score: 2046.00 Matches: 390  
Percent Similarity: 94.91% Conservatives: 20  
Best Local Similarity: 90.28% Mismatches: 22  
Query Match: 85.46% Indels: 0  
DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34287 (1-3598)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
DB 344 AAGATGGCAAGCACTTCATCCAACTTATGAAGAAAAATGAAGCTTGCGCACTGCAT 403  
QY 52 LysGlnValLeuLeuGlyProTyrAenProAspThrSerProGluValGlyPhePheAsp 71  
DB 404 AAGCAAGTTCTTATGGCCCAATATAATCCAGACACTTGTCTGAGGTTGGATTCTTTTGTAT 463  
QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
DB 464 GTGTTGGGGAATGACAGGAGGAGAGAGAAATGGCAGCCCTGGGAAACATGCTTAAGAGGAT 523  
QY 92 AlaMetValGluPheValLysLeuLeuAenLysCysProLeuLeuSerAlaTyrVal 111  
DB 524 GCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTCATCTCTTTTCAACATATGTT 583  
QY 112 AlaSerHisArgIleGluLysGluGluGluGluLysArgLysArgLysArgLysArg 131  
DB 584 GCGTCCCAAAATAGAGAAGGAAGAGAGAAAGGAAAAAGGAGGAGGAGGAGGAGGAGGAG 643  
QY 132 ArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 151  
DB 644 AGCGCGGTGNAG 703  
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluGluGlu 171  
DB 704 GAAGAAG 763  
QY 172 LeuArgLeuGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 191  
DB 764 CTTCCGTTGGAGCAGCAAG 823  
QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnGlnGln 211  
DB 824 CAGTTCAGCAGTATGACGCCCAACAGTATCCAGGGAACTACGAACAGCAGCAATTC 883  
QY 212 IleArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 231  
DB 884 ATCCGCGAGTTGAGGAGCAACACTATCAGCAGTATCAGCAGTATCAGCAGTATCAGCAG 943  
QY 232 ProAlaGlnGlnAlaAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 251  
DB 944 CTTGCAAGCAACAGGAGCACTTACAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1003  
QY 252 LeuProAlaSerSerLysValAenThrAlaGlyAlaSerAspThrLeuSerValAenGly 271  
DB 1004 TTGCTTACATCATCAAGAGTGAATGCACTGTACCAAGTAAATATGATGTCAGTTAATGGA 1063  
QY 272 GlnAlaLysThrHisThrGluAenSerGluLysValLeuGluProGluAlaAlaGluGlu 291  
DB 1064 CAGGCCAAAACACACTGACAGCTCCGAAAAAGAACTGGAAACAGAGAGAGAGAGAGAGAGAG 1123  
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311  
DB 1124 GCCTCGAAGATGGACCAAAAGAAATCTCTCCAGTAATAGCAGCTCCATCCATGTGGACA 1183  
QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331  
DB 1184 CGACCTCAGATCAAGACATTCAAAGAGAGAGATTCAGCAGGATTCAGATTCCTCGGTGATTACA 1243

QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351  
DB 1244 GTGGCCCGAGGAGAGAGTGTCTACTGTTGAGTACCACCCATGAAGAAGGATCATATCTC 1303  
QY 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371  
DB 1304 TTTTGGGAATTTGCCACAGACAAATATGACATTGGGTTGGGTTGGTATTTGAATGGACA 1363  
QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391  
DB 1364 GACTCTCCAAACACTGCTGTGAGCGTGCATGTGTCAGTGTGAGTGTGAGTGTGAGTGTGAG 1423  
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411  
DB 1424 GAAGAAGAAACATCGGTTGTGAAGAGAAAGCCAAAGAAATGCCCAACAGGCTTTGCTG 1483  
QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431  
DB 1484 GATGAGATTGTGCTGTGTACCGAGGAGTGTCTATGAGGAGTGTATGCTGGCAGCCAT 1543  
QY 432 GlnTyrProGlyArgGlyValTyrLeuLysPheAspAsnSerTyrSerLeuTrpArg 451  
DB 1544 CAATATCCAGGAGAGAGGATCTATCTCTCAAGTTTGACAACTCTACTCTTTGTGGCG 1603  
QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463  
DB 1604 TCANAATCAGTCTACTACAGAGTCTATTATATACATAGA 1639  
RESULT 5  
ADS34285  
ID ADS34285 standard; DNA; 2140 BP.  
XX ADS34285;  
AC  
DT  
XX  
XX  
DE POSH protein associated DNA #39.  
XX  
KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;  
KW antiviral; neuroleptic; central nervous system;  
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;  
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.  
XX Homo sapiens.  
XX  
PN WO2004078130-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 02-MAR-2004; 2004WO-US006308.  
XX  
PR 03-MAR-2003; 2003US-0451437P.  
PR 05-MAR-2003; 2003US-0452284P.  
PR 19-MAR-2003; 2003US-0455760P.  
PR 20-MAR-2003; 2003US-0456640P.  
PR 03-APR-2003; 2003US-0460526P.  
PR 04-APR-2003; 2003US-0460792P.  
PR 21-APR-2003; 2003US-046285P.  
PR 09-MAY-2003; 2003US-0469462P.  
PR 15-MAY-2003; 2003US-0471378P.  
PR 20-MAY-2003; 2003US-0472327P.  
PR 30-MAY-2003; 2003US-0474706P.  
PR 03-JUN-2003; 2003US-0475825P.  
PR 17-JUN-2003; 2003US-0479317P.  
PR 19-JUN-2003; 2003US-0480376P.  
PR 08-AUG-2003; 2003US-0493860P.  
PR 28-AUG-2003; 2003US-0498634P.

PR 16-SEP-2003; 2003US-0503931P.  
 PR 10-NOV-2003; 2003WO-US0035712.  
 PR 05-FEB-2004; 2004WO-US003600.  
 PR 02-MAR-2004; 2004US-0549896P.  
 XX (PROT-) PROTEOLOGICS INC.  
 PA  
 PI Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
 PI Greener T;  
 DR WPI; 2004-662346/64.  
 XX  
 PT Isolated, purified or recombinant complex, useful for identifying an  
 PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
 PT POSH-associated protein (POSH-AP).  
 XX  
 PS Disclosure; SEQ ID NO 49; 374pp; English.  
 XX  
 CC The invention relates to an isolated, purified or recombinant complex (I)  
 CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
 CC or HRPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are  
 CC useful for identifying an agent that modulates an activity of a POSH  
 CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
 CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
 CC of a protein through the secretory pathway, an agent that inhibits the  
 CC progression of a neurological disorder, an agent that modulates a POSH  
 CC function, an agent that modulates a HRPUD1 function. The methods can be  
 CC used for treating a viral infection, for inhibiting an activity of a POSH  
 CC -AP in a cell, for treating a POSH-associated disease in a subject. The  
 CC POSH-associated disease is viral infection, POSH-associated cancer or  
 CC POSH-associated neurological disorder. The methods are useful for  
 CC treating or preventing POSH-associated neurological disorder in a subject  
 CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
 CC nucleic acid of the invention.  
 XX  
 SQ Sequence 2140 BP; 628 A; 461 C; 587 G; 464 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2-27e-180 Length: 2140  
 Score: 2043.00 Matches: 389  
 Percent Similarity: 94.91% Conservative: 21  
 Best Local Similarity: 90.05% Mismatches: 22  
 Query Match: 85.34% Indels: 0  
 DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34285 (1-2140)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
 DB 337 AAAGATGGCAAGCAATTTTCATCACTTATGAAGAAAATTTGAAGCTTGTGGCACTGCAT 396  
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheApp 71  
 DB 397 AAGCAAGATTTCTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTGGATTTCTTTGAT 456  
 QY 72 ValLeuGlyAsnAspArgArgGluTyrAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
 DB 457 GTTGTGGGAATCACAGGAGGAGAGAAATGGGAGCCCTGGGAAACATGTCTAAAGAGGAT 516  
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111  
 DB 517 GCCATGGTGGATTTGTCAAGCTCTTAATAGGTGTGGCATCTCTTTTCAACATATGTT 576  
 QY 112 AlaserHisArgIleGluLysGluGluGluGluLysArgArgLysAlaGluGluArg 131  
 DB 577 GCGTCCCAAAAATAGAGAGGAGAGCAAGACAAAAGAAAGGAGGAGGAGGAGCGA 636  
 QY 132 ArgGluArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 151  
 DB 637 AGCGCGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 696  
 QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGluGluArg 171

697 GAAGAAAGAGGAAAGGCTTCGACGGAGGAAGGAAAGGAGCGATAGAGAAGAAAGG 756  
 QY 172 LeuArgLeuGluGluGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191  
 DB 757 CTTGGGTGGAGCAGCAAAAGGAGCAGATATGCGAGCTTTAAACTCCAGAGTGGCGTG 816  
 QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211  
 DB 817 CAGTTCCAGCAGTATGCAGGCCCAACAGTATCCAGGGAACCTACGACAGCAGCAATTC 876  
 QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231  
 DB 877 ATCCGCGAGTTGCGAGGAGCAACACTATCAGCAGTACATGCAGCAGTTGTATCAAGTC 936  
 QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251  
 DB 937 CTTGGCAGCAACAGGCGAGCATTTACAGAAACAACAGGAAGTAGTAGTGGCTGGCTTC 996  
 QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271  
 DB 997 TTGCTTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTCAGTTAATGA 1056  
 QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291  
 DB 1057 CAGGCCAAAACACACACTGCAGCTCCGAAAAGAACTGGAACCAAGCTGCAGAGAA 1116  
 QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311  
 DB 1117 GCCTCGAAGTGGACCAAAAGAAATCTTCTCCAGTAATAGCAGCTCCATCATGTGGACA 1176  
 QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331  
 DB 1177 CGACCTCAGATCAAAGACTTCAAAGAGAGAGATTACAGAGATGCAGATCCCGTATTA 1236  
 QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351  
 DB 1237 GTGGCGGAGGAGAGTGGTCACTGTTGAGTAGTACCACCCATGAAGAGGATCATATCTC 1296  
 QY 352 PheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThr 371  
 DB 1297 TTTTGGGAATTTGCCACAGACAAATATGACATTTGGGTTTGGGGTGTATTTTGAATG 1356  
 QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391  
 DB 1357 GACTCTCCAAACACTGCTGTGAGGTGCATGTGATGATGATGATGATGATGATGATGAT 1416  
 QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaLysProLeuLeu 411  
 DB 1417 GAAGAAAGAAACATCGTTGTGAAGAGAAAGCCAAAGAAAGTCCCAACAGCCTTTGCTG 1476  
 QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431  
 DB 1477 GATGAGATTGCTGTGTACCGCGGACTGTGTATGAGAGGAGGTGTATGTCGACGCCAT 1536  
 QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTyrArg 451  
 DB 1537 CAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGAACAACCTCTTCTTGTGGCGG 1596  
 QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463  
 DB 1597 TCAAAATCAGTCTACTACAGATCTATTATTACTAGA 1632

RESULT 6  
 AAK52051  
 ID AAK52051 standard; cDNA; 3399 BP.  
 XX  
 AC AAK52051;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 596.  
 XX





Qy 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463  
 Db 1561 TCAAAATCAGTCTACTACAGAGTCTATTATATACATAGA 1596

## RESULT 7

AD534291

ID AD534291 standard; DNA; 1995 BP.

XX

AC AD534291;

XX

DT 02-DEC-2004 (first entry)

XX

DE POSH protein associated DNA #45.

XX

KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;  
 KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
 KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI1;  
 KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
 KW anti-cancer agent; secretory pathway trafficking inhibitor;  
 KW neurological disorder progression disorder; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease; schizophrenia;  
 KW Niemann-Pick's disease.

XX

OS Homo sapiens.

XX

PN WO2004078130-A2.

XX

PD 16-SEP-2004.

XX

PF 02-MAR-2004; 2004WO-US006308.

XX

PR 03-MAR-2003; 2003US-0451437P.

PR

PR 05-MAR-2003; 2003US-0452284P.

PR

PR 19-MAR-2003; 2003US-0455760P.

PR

PR 20-MAR-2003; 2003US-0456640P.

PR

PR 03-APR-2003; 2003US-0460526P.

PR

PR 04-APR-2003; 2003US-0460792P.

PR

PR 21-APR-2003; 2003US-0464285P.

PR

PR 09-MAY-2003; 2003US-0469462P.

PR

PR 15-MAY-2003; 2003US-0471378P.

PR

PR 20-MAY-2003; 2003US-0472327P.

PR

PR 30-MAY-2003; 2003US-0474706P.

PR

PR 03-JUN-2003; 2003US-0475825P.

PR

PR 17-JUN-2003; 2003US-0479317P.

PR

PR 19-JUN-2003; 2003US-0480215P.

PR

PR 19-JUN-2003; 2003US-0480376P.

PR

PR 08-AUG-2003; 2003US-0493860P.

PR

PR 28-AUG-2003; 2003US-0498634P.

PR

PR 16-SEP-2003; 2003US-0503931P.

PR

PR 10-NOV-2003; 2003WO-US035712.

PR

PR 05-FEB-2004; 2004WO-US003600.

PR

PR 02-MAR-2004; 2004US-0549896P.

XX

PA (PROT-) PROTEOLOGICS INC.

XX

PI Taglicht DN," Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;

PI Greener T;

PI WPI; 2004-662346/64.

DR

XX Isolated, purified or recombinant complex, useful for identifying an

PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and

PT POSH-associated protein (POSH-AP).

PS Disclosure; SEQ ID NO 55; 374pp; English.

XX

CC The invention relates to an isolated, purified or recombinant complex (I)

CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)

CC or HERPUDI1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are

CC useful for identifying an agent that modulates an activity of a POSH

CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-

CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking

CC of a protein through the secretory pathway, an agent that inhibits the

CC progression of a neurological disorder, an agent that modulates a POSH

CC function, an agent that modulates a HERPUDI1 function. The methods can be

CC used for treating a viral infection, for inhibiting an activity of a POSH

CC -AP in a cell, for treating a POSH-associated disease in a subject. The

CC POSH-associated disease is viral infection, POSH-associated cancer or

CC POSH-associated neurological disorder. The methods are useful for

CC treating or preventing POSH-associated neurological disorder in a subject

CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,

CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a

CC nucleic acid of the invention.

XX

SQ Sequence 1995 BP; 628 A; 412 C; 549 G; 406 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.77e-179 Length: 1995  
 Score: 2033.00 Matches: 390  
 Percent Similarity: 94.69% Conservative: 20  
 Best Local Similarity: 90.07% Mismatches: 22  
 Query Match: 84.92% Indels: 1  
 DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34291 (1-1995)

Qy 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
 Db 314 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCAT 373  
 Qy 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71  
 Db 374 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCGAGGTGGATCTTTGAT 433  
 Qy 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
 Db 434 GTCTTGGGGAATCACAGGAGGAGAGATGGCAGCCCTGGGAAACATGCTTAAAGAGCAT 493  
 Qy 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111  
 Db 494 GCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTCCTCTCTTTTCAACATATGTT 553  
 Qy 112 AlaSerHisArgTleGluLysGluGluGluLysArgArgLysAlaGluGluAr 131  
 Db 554 GCGTCCCAAAAATAGAGAGGAAAGCAAGAAAAAAGGAGGAGGAGGAGGAGG 613  
 Qy 131 GArgGlnArgGluGluGluGluArgGluGluGlnLysGluGluLysArgLysAr 151  
 Db 614 AAGCGCGGTGAAG 673  
 Qy 151 gGluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgGluGluGluAr 171  
 Db 674 AGAAGAAGAGAGAAAGGCTTCGACGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 733  
 Qy 171 gLeuArgLeuGluGlnLysGlnLysGlnLysMetAlaAlaLeuAsnSerGlnThrAlaVa 191  
 Db 734 GCTTCGTTGGAGCAGCAAAAGCAGAGATATGGCAGCTTTTAAATCCACAGACTGCCGT 793  
 Qy 191 lGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnLe 211  
 Db 794 GCAGTTCAGCAGTATGCGAGCCCAACAGTATCCAGGGAACTACGAAACAGAGCAATTC 853  
 Qy 211 ulLeArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGl 231  
 Db 854 CATCCGCCAGTTGCGAGGCAACACATATCAGCAGTACATGCGACGAGTGTGTATCAAGTCCA 913  
 Qy 231 nProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnGlnValValMetAlaGlyAlaSe 251  
 Db 914 GCTTGCACAGCAACAGCGCAGCATTTACAGAAACAACAGGAAGTAGTAGTGGCTGCTTC 973  
 Qy 251 rLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAanGl 271  
 Db 974 CTTGCCCTACATCATCAAAAGTGAATGCAACTGTACCAGAGTAAATATGATGTCAGTTAATGG 1033  
 Qy 271 yGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGl 291

Db 1034 ACAGGCCAAACACACACACTGACAGCTCCGAAAGAACTGGAACACAGAGCTCCAGAGA 1093  
 QY 291 ualaleuGluuAsnGlyProIysAspSerLeuProValIleAlaAProSerMetTrpTh 311  
 Db 1094 AGCCCTGGAGAATGGACCAAAAGAAATCTCTCCAGTAATAGCAGCTCCATCCATGGGAC 1153  
 QY 311 rArgProGlnIleLeuAspPheIysGluuValIleArgGlnAspAlaAspSerValIleTh 331  
 Db 1154 AGACCTCAGATCAAAAGACTTCAAGAGAGATTCACGAGGATGCAGATTCCGTGATTAC 1213  
 QY 331 rValArgArgGlyGluuValIleThrValArgValProThrHisGluGluGlySerTyrIle 351  
 Db 1214 AGTGGCCCGAGGAGAGTGGTCACTCTTCAGTACCCACCCATGAGAGAGATCATATCT 1273  
 QY 351 uPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpTh 371  
 Db 1274 CTTTGTGGGAATTTGCCACAGACAATATGACATTGGGTTTGGGTTGATTTTGAATGGAC 1333  
 QY 371 rAspSerProAsnAlaAlaValSerValHisValSerGluuSerSerAspGluGluG3 391  
 Db 1334 AGACTCTCCAAACACTGCTGTGACGCTGTCATGTGAGTCCAGCGATGACGAGGAGGA 1393  
 QY 391 uGluGluGluAsnValThrCysGluGluIysAlaIysLysAsnAlaAsnLysProLeuLe 411  
 Db 1394 GGAAGAGAAACATCGTGTGTGAGAGAAAGCCAAAGAAATGCCAAAGCCTTTGCT 1453  
 QY 411 uAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHi 431  
 Db 1454 GGATGAGATTGTCCCTGTGTACCGACGGACTGTCTATGAGGAGGTGTATGCTGGCAGCCA 1513  
 QY 431 sGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpAr 451  
 Db 1514 TCAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACAACTCTCTCTTTGTGGCG 1573  
 QY 451 gSetLysSerValTyrTyrArgValTyrTyrThrArg 463  
 Db 1574 GTCAAAATCAGTCTACTACAGAGTCTATTATACTAGA 1610  
 RESULT 8  
 AAZ87207  
 ID AAZ87207 standard; cDNA; 1481 BP.  
 XX  
 AC AAZ87207;  
 XX  
 DT 08-MAY-2000 (first entry)  
 XX  
 DE Human NTAP cDNA clone 998868.  
 XX  
 KW Neurotransmission-associated protein; NTAP;  
 KW benzodiazepine receptor-associated protein; neurological disease;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW cerebral neoplasm; multiple sclerosis; drug screening; gene therapy;  
 KW antagonist; cancer; AIDS; asthma; Crohn's disease; osteoporosis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 146..1372  
 FT /\*tag= a  
 FT /product= "Human neurotransmission-associated protein  
 FT (NTAP) 998868"  
 XX  
 PN WO200001821-A2.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 02-JUL-1999; 99WO-US015121.  
 XX  
 PR 02-JUL-1998; 98US-0091677P.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Lal P, Tang YT, Yue H, Corley NC, Guegler KJ, Gorgone GA;

PI Baughn MR, Patterson C;  
 XX  
 DR WPI; 2000-160770/14.  
 DR P-PSDB; AAY77123.  
 XX  
 PT New human neurotransmission associated proteins, useful for treatment,  
 PT prevention and diagnosis of neurological disease, e.g. Alzheimer's  
 PT disease, and antagonists for treating cancer or immune disorders.  
 XX  
 PS Claim 7; Page 65; 67pp; English.  
 XX  
 CC Sequences AAZ87205-287210 represent cDNAs encoding six human  
 CC neurotransmission-associated proteins (NTAPs, AAY77121-Y77126). The  
 CC present sequence, clone 998868, encodes a benzodiazepine receptor-  
 CC associated protein, and was produced by extension of cDNA fragments  
 CC isolated from a human kidney tumour cDNA library. The NTAPs are used for  
 CC treatment or prevention of neurological diseases (e.g., Alzheimer's,  
 CC Parkinson's or Huntington's diseases, cerebral neoplasms, or multiple  
 CC sclerosis). They can also be used to raise specific antibodies and to  
 CC screen for specific binding agents (potential agonists and antagonists).  
 CC NTAP-encoding nucleic acids are useful for recombinant production of  
 CC NTAPs, and as a source of therapeutic antagonists (antisenese, triplex-  
 CC forming or ribozyme molecules). The nucleic acids may also be used as a  
 CC source of probes and primers for diagnosis or monitoring of NTAP  
 CC expression in hybridisation/amplification tests, for chromosome mapping  
 CC and for identifying related sequences, and for gene therapy. NTAP  
 CC antagonists are used to treat and prevent a wide range of cancers and  
 CC immune disorders (e.g. AIDS, asthma, Crohn's disease, osteoporosis, viral  
 CC or other infections). NTAP antibodies are used to detect NTAPs, for  
 CC diagnosis or monitoring, as therapeutic antagonists, in competitive drug  
 CC screens, and for affinity purification of NTAPs from natural sources  
 XX  
 SQ Sequence 1481 BP; 467 A; 289 C; 403 G; 322 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,58e-178 Length: 1481  
 Score: 2021.00 Matches: 385  
 Percent Similarity: 94.21% Conservative: 22  
 Best Local Similarity: 89.12% Mismatches: 25  
 Query Match: 84.42% Indels: 0  
 DB: 3 Gaps: 0  
 US-09-762-594-7 (1-463) x AAZ87207 (1-1481)  
 QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
 Db 74 AAAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAAAATTCGACCTTGGCACTGCAT 133  
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71  
 Db 134 AAGCAAGTTCTTATGGGCCCATATTAATCCAGACACTTGTCTCGAGGTTGGATTCTTTGAT 193  
 QY 72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
 Db 194 GTGTGGGGATGACAGGAGGAGAGATGGGAGCCCTGGGAAAACATGTCTAAGAGGAT 253  
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111  
 Db 254 GCCATGTGGAGTTTGTCAAGCTCTTAAATAGGTGTGCCATCTCTTTTCAACATATGTT 313  
 QY 112 AlaSerHisArgIleGluLysGluGluLysArgArgGlyAlaGluGluArg 131  
 Db 314 GCGTCCCAAAATAGAGAGGAGCAAGCAAAAAGGACGAGGAGGAGGAGCGGA 373  
 QY 132 ArgGlnArgGluGluGluArgGluArgLeuGlnLysGluGluGluLysArgLysArg 151  
 Db 374 AGCGCGCGTGAAGAGAGAGAGAGAAACGTCTGCCAAGGAGAGAGAGAAACGTAGGAGA 433  
 QY 152 GluGluGluAspArgLeuArgGluGluGluGluArgArgArgIleGluGluArg 171  
 Db 434 GAAAGAGAGGAAAGGCTTCGACGGCGGCGAGAGGAGGAGACGATAGAGAGAAAGG 493  
 QY 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191

Db 494 CTTGGTTGGAGCAGCAAAAGCAGCATATGTCAGCTTTAAACTCCAGACTGCCGTG 553  
 Qy 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnLeu 211  
 Db 554 CAGTTCAGCAGTATGAGCCCAACAGTATCCAGGAACTACGACAGCAGCAAAATCTC 613  
 Qy 212 ILAArgGlnLeuGlnGlnHisTyrGlnGlnTyrLyHisGlnAlaGluGlnThrGln 231  
 Db 614 ATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCGAGCAGTGTATCAAGTCCAG 673  
 Qy 232 ProAlaGlnGlnAlaAlaLeuGlnGlnGlnGlnGlnValMetAlaGlyAlaSer 251  
 Db 674 CTTGCACAGCAACAGGAGCATTACAGAAACAACAGGAAGTAGTGGTGGTCTTC 733  
 Qy 252 LeuProAlaSerSerLyHisValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271  
 Db 734 TTGCCTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTCAGTTATGGA 793  
 Qy 272 GlnAlaLySerThrHisThrGluAsnSerGluLyHisValLeuGluProGluAlaAlaGlu 291  
 Db 794 CAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCCAGAGCTGCAGAGAA 853  
 Qy 292 AlaLeuGlnAsnGlyProLyAspSerLeuProValIleAlaAlaProSerMetTrpThr 311  
 Db 854 GCCTCGAGAGTGGACCAAAAGATCTCTTCCAGTAATAGCAGCTCCATCCATGTGACA 913  
 Qy 312 ArgProGlnIleLyAspPheLyHisGluLyHisValSerGluSerSerValIleThr 331  
 Db 914 CGACCTCAGATCAAGACTTCAAGAGAGATTCAGCAGGATGCAATCCGGTATGACA 973  
 Qy 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351  
 Db 974 GTGGGCGAGGAGAGTGGTCACTGTTCCAGTACCCACCCATCAAGAGGATCATATCTC 1033  
 Qy 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371  
 Db 1034 TTTTGGGAATTTGCCACAGCAATATGACATTTGGGGTTTGGGGTGTATTTTGAATGACA 1093  
 Qy 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerSerSerGluGluGlu 391  
 Db 1094 GACTCTCCAAACACTGCTGTACGGTGCATGTGATGAGTCCAGGATGATGACGAGGAG 1153  
 Qy 392 GluGluGluAsnValThrCysGluGluLyHisAlaLyLyAsnAlaAsnLyProLeuLeu 411  
 Db 1154 GAAGAGAAACATCGTGTGAGAGAGAGCCAAAAGATGCCCAACAGCTTTGCTG 1213  
 Qy 412 AspGluLeuValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431  
 Db 1214 GATGAGATTGTGCTGTGTACCGAGGAGTGTATGATGAGGAGTGTATGTCGAGCCAT 1273  
 Qy 432 GlnTyrProGlyArgGlyValTyrLeuLyPheAspAsnSerTyrSerLeuTrpArg 451  
 Db 1274 CAATATCCAGGAGAGAGTGTATCTCCTCAAGTTTGACAACTCTACTCTTTGTGGCGG 1333  
 Qy 452 SerLySerValTyrTyrArgValTyrTyrArg 463  
 Db 1334 TCAAAATCAGTCTACTACAGAGTCTATTACTAGTA 1369

## RESULT 9

AAK53035

ID AAK53035 standard; cDNA; 2350 BP.

XX

AC AAK53035;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 2564.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; ss.

XX

OS Homo sapiens.

XX

PN WO200157190-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US004098.

XX

PR 03-FEB-2000; 2000US-00496914.

XX

PR 27-APR-2000; 2000US-00560875.

PR

PR 20-JUN-2000; 2000US-00598075.

PR

PR 19-JUL-2000; 2000US-00620325.

PR

PR 01-SEP-2000; 2000US-00654936.

PR

PR 15-SEP-2000; 2000US-00663561.

PR

PR 20-OCT-2000; 2000US-00693325.

PR

PR 30-NOV-2000; 2000US-00728422.

XX

(HYSE-) HYSEQ INC.

XX

Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
 Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

WPI; 2001-476283/51.

DR

P-PSDB; AAM79902.

XX

Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 in diagnosis and gene therapy.

Claim 1; Page 4841-4842; 6221pp; English.

XX

The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 cytokine, cell proliferation or cell differentiation or which may induce  
 production of other cytokines in other cell populations. The  
 polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 peptide therapy. The polypeptides have various cytokine-like activities,  
 e.g. stem cell growth factor activity, haematopoiesis regulating  
 activity, tissue growth factor activity, immunomodulatory activity and  
 activin/inhibin activity and may be useful in the diagnosis and/or  
 treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 sequence listing were missing at the time of publication

XX

SQ Sequence 2350 BP; 687 A; 497 C; 628 G; 538 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1-53e-176 Length: 2350  
 Score: 2002.50 Matches: 385  
 Percent Similarity: 93.32% Conservative: 20  
 Best Local Similarity: 88.71% Mismatches: 26  
 Query Match: 83.65% Indels: 3  
 DB: 4 Gaps: 2

US-09-762-594-7 (1-463) x AAK53035 (1-2350)

Qy

32 LysAspGlyLeuAlaPheHisProThrTyrGluGluLeuLysPheValAlaLeuHis 51

Db

339 AAGATGGCAAGCATTTTCATCACTTATGAAGAAAAATTGAAGCTTGTGGCAGTCAT 398

Qy

52 LysGlnValLeuGlyProTyrAsnProAspThrSerProGluValGlyPheAsp 71

Db

399 AAGCAAGTTCTTATGGGCCCATATATCCAGACACTTGTCTCGAGGTGGATTCCTTGTAT 458

Qy

72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyVaenMetSerLySLeuAsp 91

Db

459 GTGTGGGAATGACAGGAGAGAGATGGGAGCCCTGGGAAACATGTCTTAAGAGGAT 518

Qy

92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111

Db

519 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGGTGTGCCATCTCTTTTCAACATATGTT 578

QY 112 AlaSerHisArgIleGluLysGluGluGluGluLysArgArgIleGluGluGluArg 131  
DB 579 GCGTCCCAAAAATAGAGAAGGAAGCAAGCAAGAAAAAGGAAGGAAGGAGCGA 638  
QY 132 ArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 151  
DB 639 AGCGCGGTGAGAGGAAGGAAGAGAACGCTCTGCAAAAGAGGAAGAAACGTAGGAG 698  
QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluGluGluGluGluGluGlu 171  
DB 699 GAAGAAGAGCAAGAGCTTCGACGGGAGGAAGGAAGGAGGAGCGATAGCAAGAAAG 758  
QY 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaLeuAsnSerGlnThrAlaVal 191  
DB 759 CTTCCGTTGGAGCAGCAAAAGCAGCAGATAAATGGCAGCTTTAAACTCCAGACTGCCGTG 818  
QY 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211  
DB 819 CAGTTCACAGCAGTATGCACGCCCAACAGTATCCAGGNACTACGACGAGCAAAATTC 878  
QY 212 IleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231  
DB 879 ATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCAGCAGCTGTATCAAGTCCAG 938  
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251  
DB 939 CTTGCACACCAACAGCAGCATTTACAGAAACAAACAGGAAGTAGTAGTGGCTGGCTTCC 998  
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeu-----SerVal 269  
DB 999 TTGCTCTACATCATCAAAAGTGAA---TCCAAGTGTACCAAGTAATATGATGTCAGTTT 1055  
QY 270 AsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAla 289  
DB 1056 AATAGACAGCCCAAAACACACTGACAGCTCCGAAAAGAACTGGAAACAGAGCTGCA 1115  
QY 290 GluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMet 309  
DB 1116 GAAGAAGCCCTGGAGAAATGGACCAAAAGAAATCTCTTCCAGTAAATAGCAGCTCCATCCAG 1175  
QY 310 TrpThrArgProGlnIleLysAspPheLysGluLysIleArgGluAspAlaAspSerVal 329  
DB 1176 TGGACACGACCTCAGATCAAAGACTTCAAAGAGAGATTCAGCAGATCCAGATTCGCGT 1235  
QY 330 IleThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySer 349  
DB 1236 ATTACAGTGGCGGAGGAGAGTGGTCACTGTTCCGAGTACCACCCCATCAAGAGGATCA 1295  
QY 350 TyrLeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGlu 369  
DB 1296 TATCTCTTTTGGGAATTTGCCACAGACAATATGACATTTGGGTGGGTGTATTGTGAA 1355  
QY 370 TrpThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGlu 389  
DB 1356 TGGACAGACTCTCCAAACACTGCTGTCACGCTGCATGTCAGTGCAGTCCAGCGATGACGAC 1415  
QY 390 GluGluGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaLysPhePro 409  
DB 1416 GAGGAGGAAGAAAGAACATCGTGTGGAAGAAAGCCAAAGAAATGCCAACCAAGCT 1475  
QY 410 LeuLeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGly 429  
DB 1476 TTGCTGGATGAGATTGTGCTGTGTACCCAGCGGACTGTCAATGAGAGGTGTATGTGTCG 1535  
QY 430 SerHisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAsnSerTyrSerLeu 449  
DB 1536 AGCCATCAATATCCAGGGAGAGAGTCTATCTCTCAAGTTTGACAACTCCACTCTTGTG 1595  
QY 450 TrpArgSerLysSerValTyrTyrArgValTyrTyrThrArg 463  
DB 1596 TGGCGGTCAAAATCAGTCTACTACAGAGTCTATTACTAGTA 1637

## RESULT 10

ADS34289  
ID ADS34289 standard; DNA; 3572 BP.  
XX  
AC ADS34289;  
XX 02-DEC-2004 (first entry)  
DT  
XX POSH protein associated DNA #43.  
DE  
XX ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;  
KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;  
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.  
XX Homo sapiens.  
OS  
XX WO2004078130-A2.  
XX PD 16-SEP-2004.  
XX  
XX 02-MAR-2004; 2004WO-US006308.  
XX 03-MAR-2003; 2003US-0451437P.  
PR 05-MAR-2003; 2003US-0452284P.  
PR 19-MAR-2003; 2003US-0455760P.  
PR 20-MAR-2003; 2003US-0456640P.  
PR 03-APR-2003; 2003US-0460528P.  
PR 04-APR-2003; 2003US-0460792P.  
PR 21-APR-2003; 2003US-0464285P.  
PR 09-MAY-2003; 2003US-0469462P.  
PR 15-MAY-2003; 2003US-0471378P.  
PR 20-MAY-2003; 2003US-0472327P.  
PR 30-MAY-2003; 2003US-0474706P.  
PR 03-JUN-2003; 2003US-0475825P.  
PR 17-JUN-2003; 2003US-0479317P.  
PR 19-JUN-2003; 2003US-0480215P.  
PR 19-JUN-2003; 2003US-0480376P.  
PR 08-AUG-2003; 2003US-0493860P.  
PR 28-AUG-2003; 2003US-0498634P.  
PR 16-SEP-2003; 2003US-0503931P.  
PR 10-NOV-2003; 2003WO-US035712.  
PR 05-FEB-2004; 2004WO-US003600.  
PR 02-MAR-2004; 2004US-0549896P.  
XX (PROT-) PROTEOLOGICS INC.  
PA Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
PI Greener T;  
XX WPI; 2004-662346/64.  
XX Isolated, purified or recombinant complex, useful for identifying an  
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
PT POSH-associated protein (POSH-AP).  
XX Disclosure; SEQ ID NO 53; 374pp; English.  
XX The invention relates to an isolated, purified or recombinant complex (I)  
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
CC or HERPUD1 and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are  
CC useful for identifying an agent that modulates an activity of a POSH  
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
CC of a protein through the secretory pathway, an agent that inhibits the  
CC progression of a neurological disorder, an agent that modulates a POSH  
CC function, an agent that modulates a HERPUD1 function. The methods can be  
CC used for treating a viral infection, for inhibiting an activity of a POSH  
CC -AP in a cell, for treating a POSH-associated disease in a subject. The

CC POSH-associated disease is viral infection, POSH-associated cancer or  
 CC POSH-associated neurological disorder. The methods are useful for  
 CC treating or preventing POSH-associated neurological disorder in a subject  
 CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
 CC nucleic acid of the invention.

SQ Sequence 3572 BP; 1067 A; 665 C; 826 G; 1013 T; 0 U; 1 Other;

#### Alignment Scores:

Pred. No.: 1-74e-174 Length: 3572  
 Score: 1983.00 Matches: 385  
 Percent Similarity: 94.00% Conservative: 22  
 Best Local Similarity: 88.91% Mismatches: 25  
 Query Match: 82.83% Indels: 2  
 DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34289 (1-3572)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluLysLeuLysPheValAlaLeuHis 51  
 DB 315 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAATTGAAGCTGTGGCACTGCAT 374  
 QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheApp 71  
 DB 375 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTCAGGCTTGATTTCTTTGAT 434  
 QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
 DB 435 GTGTTGGGAATACAGAGGAGGAGAAATGGGAGCCCTGGGAAACATGCTTAAGAGGAT 494  
 QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111  
 DB 495 GCCATGGTGGATTTGTCAAGCTCTTAATAGTGTGGCATCTCTTTCAACATATGTT 554  
 QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131  
 DB 555 GCCTCCCAAAATAGAGGAAGAGCAAGACAAAGAAAGGAGGAGGAGGAGGAGGAGGAG 614  
 QY 132 ArgGlnArgGluGluGluArgGluArgGluLysGluLysGluLysArg 151  
 DB 615 AGCGCGGTGAAGAGGAAGAGAGAGCGCTCTCAAAAGAGGAGGAGGAGGAGGAGGAGGAG 674  
 QY 152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgIleGluGluArg 171  
 DB 675 GAAGAGAGAAAGCTTCGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 734  
 QY 172 LeuArgLeuGluGlnGlnLysGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191  
 DB 735 CTTCCGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAACTCCCACTCCCGCTG 794  
 QY 192 GlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGlnGlnGlnIleLeu 211  
 DB 795 CAGTTCCAGCAGTATCCAGCCCAACGGTATCCAGGGAACCTACGAAACAGCAGCAAAATTC 854  
 QY 212 IleArgGlnLeuGlnGlnIleHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231  
 DB 855 ATCCGCGCAGTTGCAGAGGACCAACTATCAGCAGTATCAGCAGGAGTTGTATCAAGTCCAG 914  
 QY 232 ProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnGlnValValMetAlaGlyAlaSer 251  
 DB 915 CTTGCACAGCAACAGCAGCAGATTACAGAAACACAGCAAGTAGTAGTGGCTGGCTTCC 974  
 QY 252 LeuProAlaSerSerLysValIleThrAlaGlyAlaSerAspThrLeuSerValIleGly 271  
 DB 975 TTGCTTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGCGCAGTTAATGGA 1034  
 QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291  
 DB 1035 CAGGCCAAACACACTGACAGCTCCGAAAGAAAGAACTGGAACCAAGCTGAGAGGAA 1094  
 QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311

DB 1095 GCCTGGAGAGTGGACCAAAAGAAATCTCTCCAGTAAATAGCAGCTCCATCCATGTGGACA 1154  
 QY 312 ArgProGlnIleLysAsp-PheLysGluLysValIleArgGlnAspAlaAspSerValIleTh 331  
 DB 1155 CGACCTCAGATCAAGACATTTCAAGAGAGAGATTACAGAGGATTCAGATTCCGTGATTAC 1214  
 QY 331 rValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLe 351  
 DB 1215 AGTGGC-CGAGGAGAGAGTGGTCACTGTTCAGTACCCACCCATGAAGAGGATCATATCT 1273  
 QY 351 upheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpTh 371  
 DB 1274 CTTTGGGAATTTGCCACAGACAATTTGTGACATTTGGGTTTGGGGTGTATTTTGAATGGAC 1333  
 QY 371 rAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391  
 DB 1334 AGACTCTCCAAACACTGCTGTGAGCGTGATGTCAAGTTTGACAACTCTCTTTGTGGCG 1393  
 QY 391 uGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLe 411  
 DB 1394 GGAAGAGAAACATCGTGTGTGAGAGAGAGCCAAAAGAAATGCCAACAAAGCTTTTGTCT 1453  
 QY 411 uAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHi 431  
 DB 1454 GGATGAGATTGTGCTGTGTACCGAGGAGCTGTATGAGGAGGTGTATGCTGGCAGCCA 1513  
 QY 431 sGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpAr 451  
 DB 1514 TCAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACAACTCTCTTTGTGGCG 1573  
 QY 451 gSerLysSerValTyrTyrArgValTyrTyrThrArg 463  
 DB 1574 GTCAAAATCAGTCTACTACAGAGTCTATTATACTAGA 1610  
 RESULT 11  
 ADS34288  
 ID ADS34288 standard; DNA; 3049 BP.  
 XX  
 AC ADS34288;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE POSH protein associated DNA #42.  
 XX  
 KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;  
 KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
 KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;  
 KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
 KW anti-cancer agent; secretory pathway trafficking inhibitor;  
 KW neurological disorder progression disorder; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease; schizophrenia;  
 KW Niemann-Pick's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004078130-A2.  
 XX  
 PD 16-SEP-2004.  
 XX  
 PF 02-MAR-2004; 2004WO-US006308.  
 XX  
 PR 03-MAR-2003; 2003US-0451437P.  
 PR 05-MAR-2003; 2003US-0452284P.  
 PR 19-MAR-2003; 2003US-0455760P.  
 PR 20-MAR-2003; 2003US-0456640P.  
 PR 03-APR-2003; 2003US-0460526P.  
 PR 04-APR-2003; 2003US-0460792P.  
 PR 21-APR-2003; 2003US-0464285P.  
 PR 09-MAY-2003; 2003US-0469462P.  
 PR 15-MAY-2003; 2003US-0471378P.  
 PR 20-MAY-2003; 2003US-0472327P.  
 PR 30-MAY-2003; 2003US-0474706P.  
 PR 03-JUN-2003; 2003US-0475825P.

PR 17-JUN-2003; 2003US-0479317P.  
 PR 19-JUN-2003; 2003US-0480215P.  
 PR 19-JUN-2003; 2003US-0480376P.  
 PR 08-AUG-2003; 2003US-0493860P.  
 PR 28-AUG-2003; 2003US-0498634P.  
 PR 16-SEP-2003; 2003US-0503931P.  
 PR 10-NOV-2003; 2003WO-US035712.  
 PR 05-FEB-2004; 2004WO-US003600.  
 PR 02-MAR-2004; 2004US-0549896P.  
 XX (PROT-) PROTEOLOGICS INC.

XX Taglicht DN, Alroy I, Reiss V, Yaar L, Ben-Avraham D, Tuvia S;  
 PI Greener T;  
 XX WPI; 2004-662346/64.

XX Isolated, purified or recombinant complex, useful for identifying an  
 PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
 PT POSH-associated protein (POSH-AP).  
 XX Disclosure; SEQ ID NO 52; 374pp; English.

XX The invention relates to an isolated, purified or recombinant complex (I)  
 CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
 CC or HERPUD1 and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are  
 CC useful for identifying an agent that modulates an activity of a POSH  
 CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
 CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
 CC of a protein through the secretory pathway, an agent that inhibits the  
 CC progression of a neurological disorder, an agent that modulates a POSH  
 CC function, an agent that modulates a HERPUD1 function. The methods can be  
 CC used for treating a viral infection, for inhibiting an activity of a POSH  
 CC -AP in a cell, for treating a POSH-associated disease in a subject. The  
 CC POSH-associated disease is viral infection, POSH-associated cancer or  
 CC POSH-associated neurological disorder. The methods are useful for  
 CC treating or preventing POSH-associated neurological disorder in a subject  
 CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
 CC nucleic acid of the invention.

XX Sequence 3049 BP; 963 A; 522 C; 633 G; 931 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5,89e-146 Length: 3049  
 Score: 1676.00 Matches: 322  
 Percent Similarity: 94.44% Conservative: 18  
 Best Local Similarity: 89.44% Mismatches: 20  
 Query Match: 70.01% Indels: 0  
 DB: 13 Gaps: 0

US-09-762-594-7 (1-463) x ADS34288 (1-3049)

QY 104 CysProLeuSerAlaTyrValAlaSerHisArgIleGluLysGluGluLys 123  
 DB 13 TGCATCTCTTTTCAATATGTTGGCTCCACAAATATAGAGAGGAGCAAGAA 72  
 QY 124 ArgArgLysAlaGluGluArgArgGlnArgGluGluGluGluArgGluGln 143  
 DB 73 AAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 132  
 QY 144 LysGluGluGluLysArgLysArgGluGluGluLysArgLysArgGluGluGlu 163  
 DB 133 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAA 192  
 QY 164 ArgArgArgIleGluGluGluArgLysArgLysGluGluGluGluGluGluGlu 183  
 DB 193 AGGAGACGATAGAGAGAGAGAGAGGCTTCGGTTGGACAGCAAAAGCAGATATGGA 252  
 QY 184 AlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyrAlaGlnGlnTyrProGly 203  
 DB 253 GCTTTAACTCCCACTGCGGTGCACTTCCAGCAGTATGCAGCCCAACAGTATCCAGG 312

QY 204 AsnTyrGluGlnGlnGlnLeuIleLeuIleArgGlnGlnGlnGlnGlnHisTyrGlnGlnTyr 223  
 DB 313 AACTACGACAGCAGCAAAATCTCATCGCCAGTTGCGAGGACCACTATCAGCAGTAC 372  
 QY 224 LysHisGlnAlaGluGlnThrGlnProAlaGlnGlnGlnAlaAlaLeuGlnGlnGlnGln 243  
 DB 373 ATGCAGCAGTGTGTATCAAGTCCAGCTTGACACAGCAACAGGCGAGCATTTACAGAAACAACAG 432  
 QY 244 GluValValMetAlaGlyAlaSerLeuProAlaSerSerLeuValAsnThrAlaGlyAla 263  
 DB 433 GAAAGTAGTAGTGGCTGCTTCCTTGGCTCATCATCAAAAGTGAATGCAACTGTACCA 492  
 QY 264 SerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysVal 283  
 DB 493 AGTAATATGATGTCTAGTTAATGGACAGGCCAAACACACACTGACAGCTCCGAAAAGNA 552  
 QY 284 LeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyProLysAspSerLeuProVal 303  
 DB 553 CTGGAACCCAGAGCTGCAGAGAGGCCCTGGAGAAATGGACCAAAAGAAATCTCTCCAGTA 612  
 QY 304 IleAlaAlaProSerMetTyrThrArgProGlnIleLysAspPheLysGluLysIleArg 323  
 DB 613 ATAGCAGCTCCATCCATCGGACAGCAGCTCAGATCAAGACTTCAAGAGAGAAATTCAG 672  
 QY 324 GlnAspAlaAspSerValIleThrValArgArgGlyGluValValThrValArgValPro 343  
 DB 673 CAGGATGAGATTCCTGATTTACATGGCCGAGGAGAGAGTGGTCACTGTTGAGTACCC 732  
 QY 344 ThrHisGluGluGlySerTyrLeuPheThrGluPheAlaThrAspSerTyrAspIleGly 363  
 DB 733 ACCCATGAAGAGGATCATATCTCTTTGGGAATTTGCCACAGACAAATATGACATGGG 792  
 QY 364 PheGlyValTyrPheGluTyrThrAspSerProAsnAlaAlaValSerValHisValSer 383  
 DB 793 TTTGGGGTGTATTTTGAATGGACAGACTCTCCAAACACTGCTGTGAGCGGTGATGTCAGT 852  
 QY 384 GluSerSerAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 403  
 DB 853 GAGTCCAGCATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 912  
 QY 404 LysAsnAlaAsnLysProLeuLeuAspGluIleValProValTyrArgArgAspCysHis 423  
 DB 913 AAGAAATGCCAACCAAGCCTTTGCTGGATGAGATTGTGCTGTGTACCGACGGACTGTCTAT 972  
 QY 424 GluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTyrLeuLysPhe 443  
 DB 973 GAGGAGGTGTATGCTGGCAGCCATCAATATCCAGGAGGAGGAGTCTATCTCTCAAGTTT 1032  
 QY 444 AspAsnSerTyrSerLeuTyrArgSerLysSerValTyrTyrArgValTyrThrArg 463  
 DB 1033 GACAACTCTACTCTTTTGGCGGTCAAAATCAGTCTACTACAGAGTCTATTATACTAGA 1092

RESULT 12

ACA56933

ID ACA56933 standard; cDNA; 961 BP.

XX ACA56933;

XX 10-JUN-2003 (first entry)

XX Human adipocyte Selected Interacting domain, SID, cDNA #20.

XX Human; ss; gene; prey; adipocyte; SID; selected interacting domain;  
 KW anorectic; antidiabetic; protein-protein interaction; diabetes;  
 KW yeast 2-hybrid assay; metabolic disorder; obesity.

XX Homo sapiens.

XX WO200286122-A2.

XX 31-OCT-2002.

XX 14-MAR-2002; 2002WO-EP003768.





```
PF 18-FEB-2000; 2000WO-US004340.
XX
PR 19-FEB-1999; 99US-0120680P.
PR 23-APR-1999; 99US-00298733.
PR 17-AUG-1999; 99US-0149639P.
PR 23-SEP-1999; 99US-0155686P.
PR 01-OCT-1999; 99US-0157247P.
PR 29-NOV-1999; 99US-0167822P.
PR 29-NOV-1999; 99US-0167823P.
PR 15-FEB-2000; 2000US-0182711P.
XX
PA (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX
XX WPI; 2000-549267/50.
DR P-PSDB; AAB23616.
XX
XX New secreted proteins and polynucleotides encoding them, which are
PT derived from Homosapiens, useful for therapy, diagnosis, and research, as
PT well as nutritional sources or supplements.
XX
XX Claim 40; Page 262-263; 309pp; English.
XX
XX The present sequence is the coding sequence for a human secreted protein.
CC It was isolated from an adult pancreas cDNA library. The proteins and
CC coding sequences of the invention can be used in the isolation of similar
CC genes and proteins, in the elucidation of their function in vivo, and to
CC treat a number of conditions. It is possible that they may have uses as
CC nutritional supplements, as cytokine or cell proliferation factors, in
CC immune modulation, where they may be used to treat immune and autoimmune
CC diseases, as haematopoiesis regulators (treating myeloid or lymphoid cell
CC deficiencies), in the promotion of tissue growth, they may have chemokine
CC or chemotactic activity, haemostatic or thrombolytic activity, or anti-
CC inflammatory activity
XX
XX Sequence 2714 BP; 837 A; 462 C; 525 G; 890 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 2,07e-99 Length: 2714
Score: 1176.00 Matches: 224
Percent Similarity: 93.25% Conservative: 11
Best Local Similarity: 88.89% Mismatches: 17
Query Match: 49.12% Indels: 0
DB: 3 Gaps: 0
US-09-762-594-7 (1-463) x AAA93116 (1-2714)
QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db 2 ATCCGCCAGTTTCAGAGGCAACACTATCAGCATATCAGCATATCAGCATATCAAGTCCAG 61
QY 232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGlnValValMetAlaGlyAlaSer 251
Db 62 CTTGCACACCAACAGCCAGCATATACAGAAACACAGCAAGTAGTAGTGGCTGGTTTCC 121
QY 252 LeuProAlaSerSerLysValLeuThrAlaGlyAlaSerAspThrLeuSerValLeuGly 271
Db 122 TTGCTTACATCATCAAAAGTGAATGCACTGTACCAAGTAATATGATGTGAGTTAATGGA 181
QY 272 GlnAlaLysThrHisThrGluLeuSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db 182 CAGGCCAAACACACACTGACAGCTCCGAAAGAAAGAACTGGAACCAAGAGCTCAGAGAA 241
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThr 311
Db 242 GCCTGGAGATGGACCAAAAGAAATCTCTCCAGTAATAGCAGCTCCATCCATGGTGGACA 301
QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db 302 CGACCTTCAGATCAAGACTTCAAGAGAGAGATTCAGCAGGATGCAGATTCCTGGATTACA 361
QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
```

```
Db 362 GTGGCCCGAGGAGAAAGTGTCTCACTGTCGAGTACCCACCATGAAGAAGGATCATATCTC 421
QY 352 PheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThr 371
Db 422 TTTTGGAAATTTGCCACAGACAAATATGACATTGGGTTTGGGGTGTATTTTGAATGGACA 481
QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
Db 482 GACTCTCCAAACACTGCTGTGATGTCATGTCAGTGTGATGTCAGGATGACACGAGGAG 541
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
Db 542 GAAGAAGAAACATCGTTGTGAAGAGAAAGCCAAAGAAATGCCAACAAAGCTTTGCTG 601
QY 412 AspGluLeuValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
Db 602 GATGAGATTGTGCTGTGTACCGAGGAGCTGTATGAGGAGGTGTATGCTGCAGCCAT 661
QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArg 451
Db 662 CAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACACTCTCTTGTGTGGCGG 721
QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
Db 722 TCAAATCAGTCTACTACAGAGTCTATTATACTAGA 757
RESULT 14
ABL24635
ID ABL24635 standard; DNA; 1495 BP.
XX
AC ABL24635;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 25378.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
DR
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 25378; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
```



CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3817 BP; 827 A; 1006 C; 1039 G; 945 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 9,74e-60 Length: 3817  
 Score: 752.50 Matches: 183  
 Percent Similarity: 45.87% Conservative: 67  
 Best Local Similarity: 33.58% Mismatches: 158  
 Query Match: 31.43% Indels: 138  
 DB: 4 Gaps: 10

US-09-762-594-7 (1-463) x ABL24634 (1-3817)

QY 34 GlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHisLysGln 53  
 DB 2569 GGCAGGCCATCCACTTGTACGAGGACACCTCAAGCTGATCGCCTTCAAGCAGCAG 2510  
 QY 54 ValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAspValLeu 73  
 DB 2509 GCGGCGCTGGGCCCTTCAACAGAGCTCGCGCCCGCGCTGGCGTCTCGACGTGATC 2450  
 QY 74 GlyAsnAspArgArgGluTrpAlaLeuGlyAsnMetSerLysGluAspAlaMet 93  
 DB 2449 GGACGCGACCGGACGACAGCTGGCAGCTGCTCGGTGAGATAACGCGCAGCAGCGGATG 2390  
 QY 94 ValGluPheValLysLeuLeuLysCysProLeuLeuSerAlaTyrValAlaSer 113  
 DB 2389 GAGGGTTCGTGACCTGTCTGGACACAATGTGCGCGCTTCGGCGGTACATTT----- 2336  
 QY 114 HisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArgGln 133  
 DB 2335 -----GAGGCGCTGGCGCAGGAC 2318  
 QY 134 ArgGluGluGluArgGluArgLysGluGluLysArgLysArgGluGlu 153  
 DB 2317 CGGACGAGACGCTGGCAGGAGCTGCGCTCATGGAGAGAAGAAGGAGGACGCGAG 2258  
 QY 154 GluAspArgArgGluGluGluArgArgArgIleGluGluArgLeuArg 173  
 DB 2257 -----CGGAGGAGAACGCCGAGGAGCTGCTCAGGAGGGGTTACAAG 2213  
 QY 174 LeuGluGlnGlnLysGlnIleMetAlaLeuAsnSerGlnThrAlaValGlnPhe 193  
 DB 2212 GAGGAGCTGCGCGCGCGCTGCGAGGAGCACTCAACAGCAGACTTACCAGCAGTTT 2153  
 QY 193 ----- 193  
 DB 2152 AAGGTAGCCCATCGATTTTTCACACACACCGCGCCCTCTGCACCTGTTGTTGT 2093  
 QY 193 ----- 193  
 DB 2092 GCTAAATATGATGATATTAATCTGGGTGCTGGCCCTTGGCTGCTGTGTCATTGAA 2033  
 QY 193 ----- 193  
 DB 2032 AAGGGGTATGACCCAGTTCTCAGGCAATTGCAATGCGAGCTCTGTATTAACATTGCG 1973  
 QY 194 -----GlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyr 205  
 DB 1972 CTTTCTAAGCTTACACTCGTTCACAGCTGTACGCTGAGAAGCAGTTCCTCCGGAACCC 1913  
 QY 205 rGluGlnGlnIleLeuIleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLys-- 224  
 DB 1912 CGAACAGAGCGCGCTTATCCACCGAGCTGCGGAGGACATATCACAGTACATGCA 1853  
 QY 225 -----HisGlnAlaGluGlnThrGlnProAlaGlnGlnAlaLeuGlnLysGln 242  
 DB 1852 ACAGCTTTCATCTCGAGAACCCAGAACCAATCAGAAATCAGAACCCAGGAGGAGGACA 1793

QY 242 nGlnGluValVal-----MetAlaGlyAlaSerLeuProAla 255  
 DB 1792 CCAGAGGAGCAGAACACGTCCTCGGCAATAACAACACGACTCAACGACTTCCCCAACGC 1733  
 QY 255 rSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysTh 275  
 DB 1732 CATGAGGAGCTGAAGCTCGGTGAGTGCAGACACAGCAGCAGCAGCAGTCCGCGGCA 1673  
 QY 275 rHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAs 295  
 DB 1672 GCAAGACGGGCATGTAGACACACGTTGGAGGCGTGGGGCAGGAACAACACGCGCAGCA 1613  
 QY 295 nGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThrArgProGlnI 315  
 DB 1612 A---GAGTACGACGACTACGTTATGATCTGCGCGCAAGATCTGGACTCTGCGCGACAT 1556  
 QY 315 eLysAspPheLysGluLysIleArgGln---AspAlaAspSerValIleThrValArgAr 334  
 DB 1555 CGAGCAGTTTAACTAGACTGAGGTGTCGCGCGCATGGGACGCGCTCATCACCATTGGACA 1496  
 QY 334 gGlyGluValValThrVal----- 340  
 DB 1495 CGGCGACACGCTGACGCT-AGGTGCAATCAACTAATGACAAACAAGCCCATATTGAC 1437  
 QY 341 -----ArgValProThrHisGluGluGlySerTyrLeuPheTrpGl 354  
 DB 1436 CGGCACAATTTTGCAGGTCGCTGCCCAACAACATGAACGGCAAGTGTCATCTCTGGGA 1377  
 QY 354 uPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThrAspSerPr 374  
 DB 1376 GTTCGCCACGACAGCTAGCATCGGCTTCGGCATCTACTTCGAGTGGGCCCAAGCCGT 1317  
 QY 374 oAsnAlaAlaValSerValHisValSerGluSerSerAspGluGlu----- 389  
 DB 1316 GACCAACGAGGTGACCGTGCATGTGAGCGACTCCGACGAGGACGAGGATTGCGTGACGA 1257  
 QY 390 -----GluGluGluGluGluAsnValThrCysGluGluLysAlaLys 403  
 DB 1256 GGACTACTTATCCACACCGAGGACCTGGAGTCCGCTCGCTGTCGAGGCGCGGCGC 1197  
 QY 403 sLysAsn-----AlaAsnLysProLeuLeuAspGluIleValProValTyrAr 419  
 DB 1196 AGTTAACACCGACTGTGCCCCCAAGCAGCACCATTTCATATCTGACCATCTATCG 1137  
 QY 419 gArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTy 439  
 DB 1136 ACGGGAGTGCTACAACGAGGTCTACGTGGGCTCCCACTCTCTATCCGGCGAGGCGCTCA 1077  
 QY 439 rLeuLeuLysPheAspAsnSerTyrSerLeuTrpArgSerLysSerValTyrTrpArgVa 459  
 DB 1076 TTTGTGAAGTTTCAGAACAGCATCTGGCGCAATAAGACCGCTGTACTACCGCGT 1017  
 QY 459 lTyrTyrThrArg 463  
 DB 1016 CTACTACGAGCGT 1004

Search completed: August 23, 2005, 02:34:15  
 Job time : 770 secs



Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

# REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

# REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

# REFERENCE

6 (bases 1 to 3433)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

# REFERENCE

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/

# FEATURES

Location/Qualifiers  
1..3433  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:5330404D19"  
/db\_xref="taxon:10090"  
/clone="5330404D19"  
/sex="male"  
/tissue\_type="pituitary gland"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
61..1638  
/note="unnamed protein product; GOLGI RESIDENT PROTEIN GCP60 homolog [Homo sapiens] (SPTPQ9H3P7, evidence: FASTV, 88.8%id, 100%length, match=1575)  
putative"  
/codon\_start=1  
/protein\_id="BAC26928.1"  
/gi\_xref="GI:26326369"  
/translation="MAAQLNVQLVSLDGLTSLSPDSEEPGAEAGPPTPSSAPGN GLGSGQQQRSPGEAAAGAAEARRMEQHMGFGLSELYGLALRYKIDGKAHPHT YEKLLKFLVALHQVLLGPNYPTSPVGFVDLGNDRRREWAALGNMSKDAMVFEVK LNKCCPLLSAVYASHRIEKEEKRAEERERLRLOKEEKREKREKED LRREERERRIIEERLRLEQKQQLMAALNSQTAVQFQYAAQQYPGNYEQQILIRQ LQEQHYQQYMQQLYQVLAQQQAALQKQEVVMAGASLPASSKVNVTAGADTLVNGQ

# CDs

1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
2219  
2220  
2221  
2222  
2223  
2224  
2225  
2226  
2227  
2228  
2229  
2230  
2231  
2232  
2233  
2234  
2235  
2236  
2237  
2238  
2239  
2240  
2241  
2242  
2243  
2244  
2245  
2246  
2247  
2248  
2249  
2250  
2251  
2252  
2253  
2254  
2255  
2256  
2257  
2258  
2259  
2260  
2261  
2262  
2263  
2264  
2265  
2266  
2267  
2268  
2269  
2270  
2271  
2272  
2273  
2274  
2275  
2276  
2277  
2278  
2279  
2280  
2281  
2282  
2283  
2284  
2285  
2286  
2287  
2288  
2289  
2290  
2291  
2292  
2293  
2294  
2295  
2296  
2297  
2298  
2299  
2300  
2301  
2302  
2303  
2304  
2305  
2306  
2307  
2308  
2309  
2310  
2311  
2312  
2313  
2314  
2315  
2316  
2317  
2318  
2319  
2320  
2321  
2322  
2323  
2324  
2325  
2326  
2327  
2328  
2329  
2330  
2331  
2332  
2333  
2334  
2335  
2336  
2337  
2338  
2339  
2340  
2341  
2342  
2343  
2344  
2345  
2346  
2347  
2348  
2349  
2350  
2351  
2352  
2353  
2354  
2355  
2356  
2357  
2358  
2359  
2360  
2361  
2362  
2363  
2364  
2365  
2366  
2367  
2368  
2369  
2370  
2371  
2372  
2373  
2374  
2375  
2376  
2377  
2378  
2379  
2380  
2381  
2382  
2383  
2384  
2385  
2386  
2387  
2388  
2389  
2390  
2391  
2392  
2393  
2394  
2395  
2396  
2397  
2398  
2399  
2400  
2401  
2402  
2403  
2404  
2405  
2406  
2407  
2408  
2409  
2410  
2411  
2412  
2413  
2414  
2415  
2416  
2417  
2418  
2419  
2420  
2421  
2422  
2423  
2424  
2425  
2426  
2427  
2428  
2429  
2430  
2431  
2432  
2433  
2434  
2435  
2436  
2437  
2438  
2439  
2440  
2441  
2442  
2443  
2444  
2445  
2446  
2447  
2448  
2449  
2450  
2451  
2452  
2453  
2454  
2455  
2456  
2457  
2458  
2459  
2460  
2461  
2462  
2463  
2464  
2465  
2466  
2467  
2468  
2469  
2470  
2471  
2472  
2473  
2474  
2475  
2476  
2477  
2478  
2479  
2480  
2481  
2482  
2483  
2484  
2485  
2486  
2487  
2488  
2489  
2490  
2491  
2492  
2493  
2494  
2495  
2496  
2497  
2498  
2499  
2500  
2501  
2502  
2503  
2504  
2505  
2506  
2507  
2508  
2509  
2510  
2511  
2512  
2513  
2514  
2515  
2516  
2517  
2518  
2519  
2520  
2521  
2522  
2523  
2524  
2525  
2526  
2527  
2528  
2529  
2530  
2531  
2532  
2533  
2534  
2535  
2536  
2537  
2538  
2539  
2540  
2541  
2542  
2543  
2544  
2545  
2546  
2547  
2548  
2549  
2550  
2551  
2552  
2553  
2554  
2555  
2556  
2557  
2558  
2559  
2560  
2561  
2562  
2563  
2564  
2565  
2566  
2567  
2568  
2569  
2570  
2571  
2572  
2573  
2574  
2575  
2576  
2577  
2578  
2579  
2580  
2581  
2582  
2583  
2584  
2585  
2586  
2587  
2588  
2589  
2590  
2591  
2592  
2593  
2594  
2595  
2596  
2597  
2598  
2599  
2600  
2601  
2602  
2603  
2604  
2605  
2606  
2607  
2608  
2609  
2610  
2611  
2612  
2613  
2614  
2615  
2616  
2617  
2618  
2619  
2620  
2621  
2622  
2623  
2624  
2625  
2626  
2627  
2628  
2629  
2630  
2631  
2632  
2633  
2634  
2635  
2636  
2637  
2638  
2639  
2640  
2641  
2642  
2643  
2644  
2645  
2646  
2647  
2648  
2649  
2650  
2651  
2652  
2653  
2654  
2655  
2656  
2657  
2658  
2659  
2660  
2661  
2662  
2663  
2664  
2665  
2666  
2667  
2668  
2669  
2670  
2671  
2672  
2673  
2674  
2675  
2676  
2677  
2678  
2679  
2680  
2681  
2682  
2683  
2684  
2685  
2686  
2687  
2688  
2689  
2690  
2691  
2692  
2693  
2694  
2695  
2696  
2697  
2698  
2699  
2700  
2701  
2702  
2703  
2704  
2705  
2706  
2707  
2708  
2709  
2710  
2711  
2712  
2713  
2714  
2715  
2716  
2717  
2718  
2719  
2720  
2721  
2722  
2723  
2724  
2725  
2726  
2727  
2728  
2729  
2730  
2731  
2732  
2733  
2734  
2735  
2736  
2737  
2738  
2739  
2740  
2741  
2742  
2743  
2744  
2745  
2746  
2747  
2748  
2749  
2750  
2751  
2752  
2753  
2754  
2755  
2756  
2757  
2758  
2759  
2760  
2761  
2762  
2763  
2764  
2765  
2766  
2767  
2768  
2769  
2770  
2771  
2772  
2773  
2774  
2775  
2776  
2777  
2778  
2779  
2780  
2781  
2782  
2783  
2784  
2785  
2786  
2787  
2788  
2789  
2790  
2791  
2792  
2793

```

Db      1117 GAAGCCCTTGAAATGACCAAAAGACTCTCTCCAGTGATTCAGCTCCATCCATGTGG 1176
Qy      311 ThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIle 330
Db      1177 ACAAGACCACAATCAAGACTTTAAGAGAAGATTCGGCAGATGCAGATTCTGTGATT 1236
Qy      331 ThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyr 350
Db      1237 ACAGTACGTCGAGGAGAAGTCGTCACCGTCCGAGTCCGACTCATGAGGAAGATCATAC 1296
Qy      351 LeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrp 370
Db      1297 CTATTTTGGGAATTTGCCACAGACAGTTATGACATTGGGTTTGGGGTTTATTTGAATGG 1356
Qy      371 ThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 390
Db      1357 ACAGACTCTCAAAATGCTGTCAGTGTCATGTGTCAGTGATCCAGTGACGAGGAGGAG 1416
Qy      391 GluGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeu 410
Db      1417 GAGGAGGAAGAAATGTCACTTGTGAAGAAAAAGCAAAAAGCAACGCCAACAGCCTCTG 1476
Qy      411 LeuAspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySer 430
Db      1477 CTGGATCGAGATTGTACCTGTGTACCGCGGGAGTGTACAGGAGGAAGTATATGCGGCAGC 1536
Qy      431 HisGlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrp 450
Db      1537 CACAGTATCCAGGAGGGAGTCTATCTCTCAAGTTGATTAATCTCTACTCTCTGTGG 1596
Qy      451 ArgSerLysSerValTyrTyrArgValTyrTyrArg 463
Db      1597 AGGTCCAAAGTCCGTCTACTACAGAGCTATTACTACTAGA 1635

```

## RESULT 2

```

BC045533
LOCUS      Homo sapiens, Similar to golgi complex associated protein 1, 60kDa,
DEFINITION clone IMAGE:5259930, mRNA.
ACCESSION BC045533
VERSION    BC045533.1 GI:28374435
KEYWORDS   HTC.
ORGANISM   Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3560)
Strausberg,R.
Direct Submission
Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

```

## REMARK

```

COMMENT
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 106 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15826851
This clone has the following problem: retained intron.

```

## FEATURES

```

source
1..3560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5259930"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC 95"
/lab_host="DH10B"
/notes="Vector: pBluescript"

```

## ORIGIN

```

Alignment Scores:
Pred. No.:      8,09e-200      Length:      3560
Score:          2046.00        Matches:      390
Percent Similarity: 94.91%      Conservative: 20
Best Local Similarity: 90.28%    Mismatches:   22
Query Match:      85.46%        Indels:       0
DB:              3              Gaps:         0

US-09-762-594-7 (1-463) x BC045533 (1-3560)

Qy      32 LysAspGlyValAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
Db      322 AAAGATGGCAAGCATTTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCAGTCGAT 381
Qy      52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
Db      382 AAGCAAGTCTTATGCGGCCATATATATCCAGACACTTGTCTCGAGGTTGGATTCTTTGAT 441
Qy      72 ValLeuGlyAsnAspArgArgGluTrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91
Db      442 GTGTGGGGAATCACAGGAGGAGAGAAATGGCAGCCCTGGGAAACATGCTCTAAAGAGGAT 501
Qy      92 AlaMetValGluPheValLysLeuLeuAsnLysCysProLeuLeuSerAlaTyrVal 111
Db      502 GCATGTGGTGAGTTTGTCAAGCTCTTAAATAGGTGTGGCATCTCTTTTCAACATATGTT 561
Qy      112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGluArg 131
Db      562 GCGTCCCAAAATAGAGAAGAGAGCAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 621
Qy      132 ArgGlnArgGluGluGluGluArgGluArgLeuGlnLysGluGluGluLysArgLysArg 151
Db      622 AGCGCGCGTGAAGAGGAAGAAAGAGAACGCTCTGCAAAAGAGGAGGAGAGAGAGAGAGAG 681
Qy      152 GluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgIleGluGluGluArg 171
Db      682 GAAGAAGAGAGAAAGGCTTCGACGGGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 741
Qy      172 LeuArgLeuGluGlnLysGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
Db      742 CTTGCGTTGGACAGCAAAAGCAGCAGATTAATGGCAGCTTTTAACTCCACAGCTCGCGTG 801
Qy      192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211
Db      802 CAGTTCACAGCATGATGCCACCCCAACAGATATCCAGGGAAGTACGAAACAGCAGCAATCTC 861
Qy      212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db      862 ATCCGCGAGTTTCAGGAGCAACATCATCATCAGCAGTACATGCAGCAGTTGTATCAAGTCCAG 921
Qy      232 ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnGluValValMetAlaGlyAlaSer 251
Db      922 CTTGCACAGCAACAGCAGCATTTACAGAAACACAGAGAGTAGTAGTGGCTGGGTCTTCC 981
Qy      252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
Db      982 TTGCGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTGAGTTATGGA 1041
Qy      272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGlu 291
Db      1042 CAGGCGCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCGGAAGCTGCAGAGAGAA 1101

```

```
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrrpThr 311
Db 1102 GCCTCGAGAAATGGACCAAAAGAATCTCTCCAGTAATAGCAGCTCCATCCATGTGGACA 1161
QY 312 AtgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db 1162 CGACCTCAGATCMAAGACTTCAAGAGAGAGATTTCAGCAGATCCAGATTCGGTGATACA 1221
QY 332 ValArgArgGlyGluValThrValArgValProThrHisGluGluGlySerTyrLeu 351
Db 1222 GTGGCCGAGGAGAGTGTCTCTGAGTACCCACCAATGAAGAGATCATATCTC 1281
QY 352 PheTrrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrrpThr 371
Db 1282 TTTTGGGAATTTGCCACAGACAAATATGACATTTGGGTGTATTTTGAATGGACA 1341
QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
Db 1342 GACTCTCCAAACACTGCTGTACGGTGCAATGTCAGTGAGTCCAGCGATGACGACGAGGAG 1401
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAlaAsnLysProLeuLeu 411
Db 1402 GAAGAAGAAACATCGTTGTGAAGAGAAAGCCCAAAAGAAATGCCCAACAGCCTTTGCTG 1461
QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
Db 1462 GATGAGATTTGTCTGTGTACCGACGGGACTGTTCATGAGGAGGTGTATGTCGACCCAT 1521
QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTrrpArg 451
Db 1522 CAATATCCAGGGAGAGAGGTCTATCTCCCAAGTTTGACAACTCTCTACTCTTTGTGGCGG 1581
QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
Db 1582 TCAAAATCAGTCTACTACAGAGTCTATTATCTAGTA 1617
```

## RESULT 3

```
BC039575
LOCUS Homo sapiens, clone IMAGE:5725757, mRNA. linear HTC 04-MAR-2003
ACCESSION BC039575
VERSION BC039575.1 GI:25058702
KEYWORDS HTC.
SOURCE Homo sapiens (human)
```

## ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Strausberg,R.

## REFERENCE

1 (bases 1 to 1995)  
Direct Submission  
Submitted (01-NOV-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
Contact: [amadan@systemsbiology.org](mailto:amadan@systemsbiology.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 84 Row: i Column: 21  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 15826951  
This clone has the following problem: retained intron.

## FEATURES

source  
1..1995  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5725757"  
/tissue\_type="Ovary, pooled from 3 adults"  
/clone\_lib="NIH\_MGC\_125"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

## ORIGIN

Alignment Scores: 7.9e-199 Length: 1995  
Pred. No.: 2033.00 Matches: 390  
Score: 94.69% Conservative: 20  
Percent Similarity: 90.07% Mismatches: 22  
Best Local Similarity: 84.92% Indels: 1  
Query Match: 3 Gaps: 0  
DB: 3  
US-09-762-594-7 (1-463) x BC039575 (1-1995)  
QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51  
Db 314 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAAATTGAAGCTTGTGGCACTGCAT 373  
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheApp 71  
Db 374 AAGCAAGTTCTTATGGCCCATATAATCCAGACACTTGTCTCAGGTTGGATTCTTTGAT 433  
QY 72 ValLeuGlyAsnAspArgArgGluTrrpAlaAlaLeuGlyAsnMetSerLysGluAsp 91  
Db 434 GTGTGGGGAATGACAGGAGGAGAGATGGGAGCCCTGGGAAACATGTCTAAAGAGGAT 493  
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111  
Db 494 GCATGTGTGGAGTTTGTCAAGCTCTTAATAGGTGTGCCATCTCTTTTCAACATATGT 553  
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGluAr 131  
Db 554 GCGTCCCAAAATAGAGAAGCAAGAGCAAGAAAAAAGAGAGAGAGAGAGAGAGAGCG 613  
QY 131 GArgGlnArgGluGluGluGluArgGluGluGlnLysGluGluLysArgLysAr 151  
Db 614 AAGGCGCGGTGAAGAGAGAGAGAGAACGTCTGCAAAAGAGGAGAGAGAGAGAGAGAG 673  
QY 151 gGluGluGluAspArgLeuArgArgGluGluGluArgArgArgIleGluGluGluAr 171  
Db 674 AGAAGAGAGAGAGAGGCTTCGACGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 733  
QY 171 gLeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVa 191  
Db 734 GCTTCGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACCTCCAGACTGCCGT 793  
QY 191 lGlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLe 211  
Db 794 GCAGTTCACAGTATGACGCCCAACAGTATCAGGGAACCTACGAAACAGCAGCAAAATCT 853  
QY 211 uIleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGl 231  
Db 854 CATCCGCCAGTTGACGAGGAGCAACACTATCAGCAGTACATGCAGCAGCTTGTATCAAGTCCA 913  
QY 231 nProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSe 251  
Db 914 GCTTGACACAGCAACAGCAGCATTTACAGAAACACAGGAGAGTAGTAGTGGCTGGTCTTC 973  
QY 251 rLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGl 271  
Db 974 CTTGGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTAGTTAATGG 1033  
QY 271 yGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGl 291  
Db 1034 ACAGGCCAAACACACACTGACAGCTCCGAAAAAGAACTGGAAACAGAGAGCTGCAGAGA 1093







```

Db      963  TTTTGGGAATTTCCACAGCAATATTGACATTTGGGTTTGGGGTGATTTTGAATGGACA 1022
Qy      372  AspSerProAsnAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
Db      1023  GACTCTCCAAACACTGCTGTACAGGTCATGTGAGTGAGTCCAGCGATGACGACGAGGAG 1082
Qy      392  GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411
Db      1083  GAAGAAGAAACATCGGTTGTGAAGAGAAAGCCAAAGAAAGAAATCCCAACAGCCTTTGCTG 1142
Qy      412  AspGluLeuValProValThrArgArgAspCysHisGluGluValThrAlaGlySerHis 431
Db      1143  GATGAGATTGTGCTGTGTACCGAGCGGACTGTCTCATGAGGAGGTGTATGCTGGCGCCAT 1202
Qy      432  GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAspAsnSerTyrSerLeuTyrArg 451
Db      1203  CAATATCCAGGAGAGAGTCTATCTCCTCAAGTTTGACAACTCTTACTCTTTGTGGCGG 1262
Qy      452  SerLysSerValTyrTyrArgValTyrTyrThrArg 463
Db      1263  TCAAAATCAGTCTACTACAGAGTCTATTATCTACTAGA 1298

```

## RESULT 6

```

AY403084
LOCUS      1301 bp      DNA      linear      GSS 15-DEC-2003
DEFINITION Pan troglodytes GOCAP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION  AY403084
VERSION    AY403084.1 GI:39759067
KEYWORDS   Pan troglodytes (chimpanzee)
SOURCE     GSS.
            Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
            1 (bases 1 to 1301)
            Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
            Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, P., Murphy, B.,
            Ferrier, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J.,
            Adams, M.D. and Cargill, M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED    14671302
REFERENCE  2 (bases 1 to 1301)
AUTHORS    Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
            Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, P., Murphy, B.,
            Ferrier, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J.,
            Adams, M.D. and Cargill, M.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT    This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
            source          1..1301
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        <1..>1301
                        /genes="GOCAP1"
                        /locus_tag="HCM1439"
            gene
            ORIGIN
Alignment Scores:
Pred. No.:      5,41e-162      Length:      1301
Score:          1674.00      Matches:      330
Percent Similarity: 79.86%      Conservative: 15
Best Local Similarity: 76.39%      Mismatches: 87
Query Match:    69.92%      Indels:      0
DB:              9      Gaps:      0
US-09-762-594-7 (1-463) x AY403084 (1-1301)

```

```

Qy      32  LysAspGlyLysAlaPheHisProThrTyrGluLysLeuLysPheValAlaLeuHis 51
Db      3  AAAGATGGCAAGCAATTTTCATCAACTTATGAAGAAAATTTGAAGCTTGTGGCACTGCAT 62
Qy      52  LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
Db      63  AAGCAAGTTCTTTATGGGCCCATATATCCAGACACTTGTCTCGAGGTGGATTCTTTGAT 122
Qy      72  ValLeuGlyAsnAspArgArgGluTyrAlaLeuGlyAsnMetSerLysGluAsp 91
Db      123  GTGTTGGGGAATGACAGGAGGAGAAATGGGAGCCCTGGGAAACATNNNNAAGAGGAT 182
Qy      92  AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
Db      183  GCCATGNNAGTTTGTCAAGCCTTAAATAGTNTTGCNNNNNNNTTTCACATATGTT 242
Qy      112  AlaSerHisArgIleGluLysGluGluGluLysArgLysAlaGluGluGluArg 131
Db      243  GCGTCCCCCAAAATACAGAAAGGAGCAAGCAAGAAAAAAGNNNNNNNNNNNNNNNN 302
Qy      132  ArgGlnArgGluGluGluGluArgGluGluGlnLysGluGluLysArgLysArg 151
Db      303  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 362
Qy      152  GluGluGluAspArgLeuArgArgGluGluGluGluArgArgArgIleGluGluGluArg 171
Db      363  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 422
Qy      172  LeuArgLeuGluGlnLysGlnIleMetAlaLeuAsnSerGlnThrAlaVal 191
Db      423  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 482
Qy      192  GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGlnGlnGlnGlnLeu 211
Db      483  CAGTTCCAGCAGTATGCAGCCCAACAGTAGTCCAGGGAATACGAGCAGCAAAATCTC 542
Qy      212  IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db      543  ATCCGCGAGTTCCAGGAGCACTATTCAGCAGTATATCAGCAGGTGTATCATCAAGTCCA 602
Qy      232  ProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSer 251
Db      603  CTTGCACAGCAACAGCAGCAGTATACAGAAACACAGGAAGTAGTAGTGGCTGGCTTCC 662
Qy      252  LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271
Db      663  TTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAAGTAATGGA 722
Qy      272  GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaGluGlu 291
Db      723  CAGGCCAAAACACACACTGACAGCTCCGAAAAGAAAGAACTGGAAACAGAGCTGCAGAGAA 782
Qy      292  AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311
Db      783  GCCCTGGAGAAATGGACCAANNAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGACA 842
Qy      312  ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
Db      843  CGACCTCAGATCAAAAGACTTCAAAGAGNNGATTTCAGCAGAGTGCAGATTCCGTGATTA 902
Qy      332  ValArgArgGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
Db      903  GTGGCCGAGGAGAAAGTGTCTACTGTTTCAGTAGTACCACCCTCAAGAAAGATCATATCTC 962
Qy      352  PheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThr 371
Db      963  TTTTGGGAATTTGCCACAGACAATATTGACATTTGGTTGGGGTGTATTTTGAATGGACA 1022
Qy      372  AspSerProAsnAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 391
Db      1023  GACTCTCCAAACACTGCTGTCTCAGCGGTGATGTCTCAGTAGTCCAGCGATGCCAGCAGGAG 1082

```

QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLysProLeuLeu 411  
 Db 1083 GAAGAAGAAACATCGTTGTGAAGAGAAAGCCAAAGAAATGCCAACAGCCTTTGCTG 1142  
 QY 412 AspGluLeuValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431  
 Db 1143 GATGAGATTGCTGCTGTGTACCGAGCGGACTGTCTATGAGGAGGTGTATGCTGGCAGCAT 1202  
 QY 432 GlnTyrProGlyArgGlyValTyrLeuLeuLysPheAsnSerTyrSerLeuTyrArg 451  
 Db 1203 CAATATCCAGGAGAGAGAGTATCTCTCAAGTTTGACAACTCTCTACTCTTTGTGGAGG 1262  
 QY 452 SerLysSerValTyrTyrArgValTyrTyrArg 463  
 Db 1263 TCAAAATACGTCTACTACAGAGTCTATTATACAG 1298

RESULT 7  
 BUI53320  
 LOCUS BUI53320 884 bp mRNA linear EST 03-SEP-2002  
 DEFINITION AGENCOURT\_7938530 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:6012409  
 5', mRNA sequence.

ACCESSION BUI53320  
 VERSION BUI53320.1 GI:22666852  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13203 row: k column: 02  
 High quality sequence stop: 645.

FEATURES  
 source Location/Qualifiers  
 1..884  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6012409"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 92"  
 /note="Organ: Testis; Vector: pCMV-SPORT6; Site 1: Noti;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,586-123 Length: 884  
 Score: 1295.00 Matches: 254  
 Percent Similarity: 91.53% Conservative: 16  
 Best Local Similarity: 86.10% Mismatches: 24  
 Query Match: 54.09% Indels: 1  
 DB: 5 Gaps: 0  
 US-09-762-594-7 (1-463) x BUI53320 (1-884)

QY 153 GluGluAspArgLeuArgGluGluGluArgArgGluGluGluGluGluGluGluGluGlu 172  
 Db 2 GAAGAGAAAGGTTCCGACGGGAGGAAGGAAGGAGCGGATAGACGAAGAAAGGCTT 61

QY 173 ArgLeuGluGlnGlnLysGlnInileMetAlaAlaLeuAsnSerGlnThrAlaValGln 192  
 Db 62 CGGTTGGAGCAGCAAAAGCAGCAGATATATGGCAGCTTTAACTCCAGACTCCGCTGCAG 121  
 QY 193 PheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnLeuLeu 212  
 Db 122 TTCCAGCAGTATGCAGCCCAACAGATATCCAGGGAATACGAACAGCAGCAAAATCTCATC 181  
 QY 213 ArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGlnPro 232  
 Db 182 CGCCAGTTGCAGAGCAACACTATCAGCAGTATCAGCAGTATGTATCAAGTCCAGCTT 241  
 QY 233 AlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGlyAlaSerLeu 252  
 Db 242 GCACAGCAACAGCGCAGTATCAGAAACAAACAGGAAGTAGTAGTGGCTTCCTCTG 301  
 QY 253 ProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGln 272  
 Db 302 CCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATATGATGTCAGTTAATGGACAG 361  
 QY 273 AlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluAla 292  
 Db 362 GCCAAAACACACACTCAGCAGCTCCGAAAAGAACTGGAACCAAGAGCTGCAGAAGAGCC 421  
 QY 293 LeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTrpThrArg 312  
 Db 422 CTGGAGAATGGACCAAAAGATCTCTCCAGTAATAGCAGCTCCATCCATGTGGACAGA 481  
 QY 313 ProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThrVal 332  
 Db 482 CCTCAGATCAAGAGCTTCAAGAGAAAGATTCAGCAGGATGCAGATTCGTGATTACAGTG 541  
 QY 333 ArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeuPhe 352  
 Db 542 GCCCGAGGAGAGTGGTCACTGTCGAGTACCACCCATCAAGAAAGGATCATATCTCTTT 601  
 QY 353 TrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTrpThrAsp 372  
 Db 602 TGGGAATTTGCCACAGCAATATGACATTTGGGTTGGGTGTATTTTGAATGGACAGAC 661  
 QY 373 SerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGluGlu 392  
 Db 662 TCTCCAAACACTGCTGTGTCAGCGTGCATGTCTAGTGTCCAGCATGCAGCAGGAGGAA 721  
 QY 393 GluGluAsnValThrCysGluLysAlaLysLysAsnAlaAsnLysProLeuLeuAsp 412  
 Db 722 GAAGANAACATCGTTGTGAAGAGAAAGCCCAAAAGAAATGGCAACAGCCTTTGCTGGAT 781  
 QY 413 GluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHisGln 432  
 Db 782 GAGATTGTCCTGTGTACCCGAGGACTGTCTATGGAGGAGTGTATGCTGGCAGCCATCAA 841  
 QY 433 Tyr-ProGlyArgGlyValTyrLeuLysPheAspAsnSer 446  
 Db 842 TATCCAGGAGAGAGAGTCTATCTCCCTCAGTTTGACCACTCC 884

RESULT 8  
 CF240143  
 LOCUS CF240143 918 bp mRNA linear EST 05-AUG-2003  
 DEFINITION AGENCOURT\_15112395 NICHD XGC Emb6 Xenopus tropicalis cDNA clone  
 IMAGE:6991831 5', mRNA sequence.  
 ACCESSION CF240143  
 VERSION CF240143.1 GI:33443351  
 KEYWORDS EST.  
 SOURCE xenopus tropicalis (western clawed frog)  
 ORGANISM xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 REFERENCE 1 (bases 1 to 918)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)











Qy	240	GlnIysGlnGlnGluValValMetAlaGlyAlaSerLeuProAlaSerSerIysValAsn	259
Db	2	CAGAAGCAGCAAGAAGTAGTGTGGTGGCTTCATTGCCTGCGTCAGCAAGAGTGAAC	61
Qy	260	ThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaIysThrHisThrGluAsn	279
Db	62	ACAGCTGGAGCAGGTGATCCCTGCGCAGTGAACGGACAGGCCAAGACCACACTGAGAT	121
Qy	280	SerGluIysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyProIysAsp	299
Db	122	CCCAGAGAAAGTCCCTTGAGCCAGAACTCGGAAAGAGGGCTTGGAATAATGGACCCAAAGAC	181
Qy	300	SerLeuProValIleAlaalaProSerMetTrpThrArgProGlnIleIysAspPheIys	319
Db	182	TCTCTTCAGTAGATTGCTGCTCTTCATGTGGACAAGCACACAGATCAAGACTTTAA	241
Qy	320	GluIysIleArgGlnAspAlaAspSerValIleThrValArgArgGlyGluValValThr	339
Db	242	GAGAAGATTTCGCAGAGATGCAGATTCTGTGATTACAGTCCGCGAGGAGAAAGTAGTCACC	301
Qy	340	ValArgValProThrHisGluGluGlySerTyrLeuPheTrpGluPheAlaThrAspSer	359
Db	302	GTTTCGAGTCCCACATCATGAGGAAGGATCATACCTCTTTTGGGAATTTGCCACACACAGT	361
Qy	360	TyrAspIleGlyPheGlyValTyrPheGluTrpThrAspSerProAsnAlaAlaValSer	379
Db	362	TATGACATTGGGTTTGGGGTTTATTTTGAATGGACAGACTCTCCAAAACGCTGCTGTCAGT	421
Qy	380	ValHisValSerGluSerSer---AspGluGluGluGluGluGluAsnValThrCys	398
Db	422	GTGCATGTGCAGTGATGCCAGTGACACGAAGAGGAGGAGGAAGAAGAAATGCTCACTGT	481
Qy	399	GluGluIysAlaIysLysAsnAlaAsnIysProLeuLeuAspGluIleValProValTyr	418
Db	482	GAAAGAAAGCCAAAAGAAAGTCAACAAGCCCTCTGCTGGATCAGATTGTACTGTGTAC	541
Qy	419	ArgArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyVal	438
Db	542	CGGCGGGACTGTCATGAGGAAGTGTATGCAGGCAGCCACAGTACCCAGGGAGGGGGTTC	601
Qy	439	TyrLeuLeuLysPheAspAsnSerTyrSerLeuTrpArgSerLysSerValTyrTrpArg	458
Db	602	TATCTCTCAAGTTTGATTAACCTCCACTCTCTGTGTGAGGTCTTAAGTCAGCTACTACAGA	661
Qy	459	ValTyrTyrThrArg	463
Db	662	GTCTATTACTACAGA	676

RESULT 12					
CF224319					
LOCUS	CF224319	993 bp	mRNA	linear	EST 04-AUG-2003
DEFINITION	AGENCOURT_15064731	NICHD_XGC_Emb7	Xenopus tropicalis	cdna clone	
	IMAGE:6977632	5',	mRNA	sequence.	

RECORD NUMBER	SOURCE	ORGANISM	ESL
1	Xenopus tropicalis (western clawed frog)	Xenopus tropicalis	
2	Xenopus tropicalis	Xenopus tropicalis	
3	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.		

REFERENCE 1 (bases 1 to 993)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics  
National Cancer Institute / NCI  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: csgabgs-remail@nih.gov  
Tissue Procurement: Robert M. Grainger  
CDNA Library Preparation: Life Technologies  
CDNA Library Arraved by: The I.M.A.

```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14632 row: 1 column: 15
High quality sequence stop: 715.
Location/Qualifiers
1. 993
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6977632"
/tissue_type="tailbud"
/dev_stage="embryo, stages 20-27"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG XGC Emb7"
/notes="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

```

## ORIGIN

<b>Alignment Scores:</b>					
Pred. No.:	7, 67e-106	Length:	993		
Score:	1129.00	Matches:	237		
Percent Similarity:	83.48%	Conservative:	41		
Best Local Similarity:	71.17%	Mismatches:	43		
Query Match:	47.16%	Indels:	14		
DB:	7	Gaps:	4		
US-09-762-594-7 (1-463) x CF224319 (1-993)					
Qy	115	ArgIleGluLysGluGluGluYsAqAtGLysAlaGlUGluGluAqArgGlnAtg	134		
Dd	1	CAGATCGGGGAGGAGCAGGAGAGGAGGAGGAGGAGGAGGAGCTCGAGGAAA	60		
Qy	135	GluGluGluGluArgGluArgLeuGlnLysGluGluGluLysArgLysArgGluGlu	154		
Dd	61	GAGGAGGAGCGAGGAGCGCTTACAGAGAGAAGAAAGCGGAAGAGAGGAG	120		
Qy	155	AspArgLeuArgArgGluGluGluUArgArgArgIleGluGluGluArgLeuArgLeu	174		
Dd	121	GAGAGGCAGAGGCCGAGGAAGAGGAGAGGTTCGGGCCGCAAGCACGAGCGTTCCGTATG	180		
Qy	175	GluGlnGlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaValGlnPheGln	194		
Dd	181	GAGCAGAGAAGCAGCAGATTATGGCAGCCCTGAACTCCACAGCAGCATGTCCAA	240		
Qy	195	GlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeuIleArgGln	214		
Dd	241	CAGTACGGGCCCAGCAGTACCCTCCGGCACTTTGAGCAGCAGCAGATCCTTATCCGGCAG	300		
Qy	215	LeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGlnProAlaGln	234		
Dd	301	CTCCAGGAGCAGCATTTACCAGCAGTAGTCATCAACAGCTCTACCAAGTGCAACTGGCCAC	360		
Qy	235	GlnGlnAlaAlaLeuGlnLysGlnGlnGluValValMetalGlyAlaSerLeuPro---	253		
Dd	361	CAGCAGCGCGCTTATCAGAAAGCAACAGGAGGAGCAGCAGTAGCCCCGATCTCTCGCTC	420		
Qy	254	---AlaSerSerLysValAsnThrAlaGlyAlaSer-----AspThrLeuSerVal	269		
Dd	421	ACCTCCCTCCAAGATGCACTCAGTAGGCGCCGATCCCCCGGTGAGACCCCTTCAGTC	480		
Qy	270	AsnGlyGlnAlalysThrHisThrGlnAsnSerGluLysValLeuGluProGluAlaAla	289		
Dd	481	AACGGCCAGAGTCACCTCTCACTTAGACAGCAATGAACCAAGAAATGGAGTCTGACCCACTG	540		
Qy	290	GluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaLaProSerMet	309		
Dd	541	GAGATCTTCTGAGAAATGGACCAAGAGTCGCCCCCATGTAGTCGCGCACCTTCATG	600		
Qy	310	TripThrArgProGlnIleLysAspPheLysGluLysIleArqGlnAspAlaAspSerVal	329		

Db 601 TGGACACGCCCCAGATATAAGACTTCAAGAGAGATCCGGCAGGACGCCGACTCCGTG 660  
 Qy 330 IleThrValArgGlyGluValValThrValArgValProThrHisGluGluGlySer 349  
 Db 661 ATACCGTGGCGCGCGAGGTAGTGACCGTGGAGTCCGCCACCCACGAGGAGGCTCC 720  
 Qy 350 TyrLeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGlu 369  
 Db 721 TAGCTCTTCGTGGAGTTTGCACCGGATAATTATGACATTGGTTCCGGGTATTATTGAA 780  
 Qy 370 TrpThrAspSerProAsnAlaAlaValSerValHisValSerGluSerAspGluGlu 389  
 Db 781 TGGACGCACTCNCCTTATACCGCGCTGAGTGTGCATGTTCAGCGAGTCCAGTGAGGATGAG 840  
 Qy 390 GluGluGluGlu-----GluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsn 407  
 Db 841 GAGAGATGAGGGTAAAGAAACCCACACGCGAAGAGAAACCAA-AGAAGCGCAAC 898  
 Qy 408 LysProLeuLeuAspGlu-IleVal-ProValTyrArgArgAspCys-----HisGluG 425  
 Db 899 CAGCCGCACTGACGACGAAGATCGTGCCTGGTGTCCGAAAGGGATTGGCCCCCAGGAAG 958  
 Qy 425 luValTyrAlaGly-SerHisGlnTyrProGly 435  
 Db 959 GGGTTACCCCGCGCTGCCACCACTAGTACCCCGG 991

RESULT 13  
 CK960730 742 bp mRNA linear EST 15-MAR-2004  
 LOCUS 4101641 BARC 10BOV Bos taurus cDNA clone 10BOV8\_D16 5', mRNA  
 DEFINITION CK960730

ACCESSION CK960730  
 VERSION 1 GI:45475110

KEYWORDS EST.  
 SOURCE Bos taurus (cow)

ORGANISM  
 Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovinae; Bos.

REFERENCE  
 1 (bases 1 to 742)  
 Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,  
 G.P., Bosak, S., Rubinfeld, M. and Gasbarre, L.C.  
 Production of EST from cDNA libraries derived from immunologically  
 activated bovine gut  
 Unpublished (2004)

JOURNAL  
 COMMENT  
 Contact: Tad S. Sonstegard  
 Bovine Functional Genomics Laboratory  
 Animal and Natural Resources Institute  
 Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA  
 Tel: 3015048416  
 Fax: 3015048414

Email: tad@anri.barc.usda.gov  
 Single pass sequencing. Bases called and trimmed with phred  
 0.000925 using options -trim alt -trim fasta. Vector identified  
 by cross match using options -minmatch 12 -minscore 18  
 Plate: 8 row: D column: 16  
 Seq primer: CCAGTCACGACGTTGTAACG  
 High quality sequence stop: 742.

FEATURES  
 source

1. 742  
 Location/Qualifiers  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /strain="Holstein"  
 /db\_xref="taxon:9913"  
 /clone="10BOV8\_D16"  
 /sex="Male"  
 /tissue type="Pooled"  
 /dev stage="Multiple"  
 /lab\_host="DH10B tr phage resistant"  
 /lab\_lib="BARC 10BOV"  
 /note="Organ: Small Intestine; Vector: pagen-1; Site 1:  
 EcoRV; Site 2: NotI; Equimolar amounts of mRNA extracted

# ORIGIN

## Alignment Scores:

Pred. No.: 1,12e-104 Length: 742  
 Score: 1116.00 Matches: 218  
 Percent Similarity: 92.68% Conservative: 10  
 Best Local Similarity: 88.62% Mismatches: 18  
 Query Match: 46.62% Indels: 0  
 DB: 7 Gaps: 0

US-09-762-594-7 (1-463) x CK960730 (1-742)

Qy 137 GluGluArgGluArgLeuGlnLysGluGluGluLysArgLysArgGluGluGluAspArg 156  
 Db 5 GATGAACGGGAACGCTCTGCAAAAGGAGGAAGAAACGTTAGGAGAGAGAAGAAAGG 64  
 Qy 157 LeuArgArgGluGluGluGluArgArgGileGluGluGluArgLeuArgLeuGluGln 176  
 Db 65 CTTAGACGGGAGAGAGAGAGAGGCGGCTAGAGAAGAGAGACTTCGCTGGAGCAG 124  
 Qy 177 GlnLysGlnGlnIleMetAlaAlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyr 196  
 Db 125 CAAAAGCAGCAGATAATGCCAGCTTTAAACTCCAGACTGCCGTGCAGTTCACGACGAT 184  
 Qy 197 AlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeulleArgGlnLeuGln 216  
 Db 185 GCAGCCCAACAGTATCCAGGGAACACGAAACAGCAGCAGATCTCATCCGCCAGTTGCAG 244  
 Qy 217 GluGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrClnProAlaGlnGln 236  
 Db 245 GAGCAACACTATCAGCAGTATATGAGCAGTGTGTATCAAGTCCAGCTTCACACGACAC 304  
 Qy 237 AlaAlaLeuGlnLysGlnGlnGluValValMetAlaGluValAlaSerLeuProAlaSer 256  
 Db 305 GCAGCATTTACAGAAACACACAGAGAGTAGCAGTAGTGGGGCTTCCTCTCATCATCA 364  
 Qy 257 LysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHis 276  
 Db 365 AAGTGAATGCAACTACACCAAGCATATGATGTAGTAAATGGACAGGCCAAACCCAC 424  
 Qy 277 ThrGluAsnSerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGly 296  
 Db 425 ACTGACAACTGAAAAAGAGACTGAACCAAGAGCGCGAAGAGCCCTGGAGAATGGA 484  
 Qy 297 ProLysAspSerLeuProValIleAlaAlaProSerMetTyrThrArgProGlnIleLys 316  
 Db 485 CCAAAAGAAATCTCTTCCAGTATTCGAGCTCCATCCATGTGGACACGACCCAGATCAA 544  
 Qy 317 AppPheLysGluLysIleArgGlnAspAlaAspSerValIleThrValArgArgGlyGlu 336  
 Db 545 GACTTTAAAGAGAGAGATTCGGCAGGATCGGATTCGGTGATCACAGTGGGCCGAGAGAA 604  
 Qy 337 ValValThrValArgValProThrHisGluGluGlySerTyrLeuPheTrpGluPheAla 356  
 Db 605 GTAGTCACCTGTTCCGAGTACCCACCATGAAGAAGGATCATATCTCTTTTGGGAATTTG 664  
 Qy 357 ThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThrAspSerProAsnAla 376  
 Db 665 ACAGACAATTATGACATTTGGGTGGGTATATTTTGAATGGACAGACTCTCCGAACT 724  
 Qy 377 AlaValSerValHisVal 382  
 Db 725 GCTGTGCGGTGCATGTC 742

## RESULT 14

CN310226

LOCUS

DEFINITION. 17000424186898 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN310226

from proximal jejunums of 18 and 21 wk old steers, and  
 distal ileums of 14 day old calves. proximal jejunum  
 exposed to C. oncophora for 3 and 6 weeks, and distal  
 ileum exposed to C. parvum for 7 days"

```

CN310226.1 GI:47326640
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Brandenberger R., Wei, H., Zhang S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert length: 792 Std Error: 0.00.
Location/Qualifiers
1. .792
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/notes="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN
Alignment Scores:
Pred. No.: 7 03e-103 Length: 792
Score: 1099.00 Matches: 221
Percent Similarity: 91.44% Conservative: 14
Best Local Similarity: 85.99% Mismatches: 19
Query Match: 45.91% Indels: 3
DB: 7 Gaps: 0

US-09-762-594-7 (1-463) x CN310226 (1-792)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
Db 27 AAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCAT 86
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
Db 87 AAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCTGAGGTTGGATTCTTTGAT 146
QY 72 ValLeuGlyAsnAspArgArgGluTTPAlaAlaLeuGlyAsnMetSerLysGluAsp 91
Db 147 GTGTTGGGAGATGACAGGAGGAGAGATGGCGACCCCTGGGAAACATGTCTAAAGAGGAT 206
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerLafyrVal 111
Db 207 GCCATGGTGGAGTTTGTCAAGCTCTTAATAGGTGTGCCATCTCTCTTTCAACATATGTT 266
QY 112 AlaSerHisArgIleGluLysGluGluGluGlu--LysArgArgLysAlaGluGluGlu 131
Db 267 GCGTCCCAAAAATAGAGAAGGAAGCAAGCAAGAAAAAAGGAAGGAGGAGGAGGAGC 326
QY 131 rGArgGlnArgGluGluGluGluArgGluArgLeuGlnLysGluGluLysArgLysA 151
Db 327 GAAGCGCGGTGAAGAGGAGAGAAAGAGACCTCTGCAAAAGGAGGAGGAACGTAGGA 386
QY 151 rGluGluGluAspArgLeuArgGluGluGluGluArgArgArgGluGluGluGlu 171
Db 387 GACAAGAGAGGAAAGGCTTCGACGGGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAG 446

```

```

QY 171 rGLeuArgLeuGluGlnGlnLysGlnGlnMetAlaLeuAsnSerGlnThrAlaVal 191
Db 447 GGCCTTCGGTTGGAGCAGCAAAAGCAGCAGATATATGGCAGCTTTAAACTCCAGACTGCG 506
QY 191 aGlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleL 211
Db 507 TGCAGTTCAGCAGTATGCAGCCCAACAGTATCCAGGGAACCTACGAAACAGCAGCAAAATTC 566
QY 211 euLlAArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrG 231
Db 567 TCATCCGCGCAGTTCGAGGAGCAACTATCAGCAGTACATGCAGCAGTTGTATCAAGTCC 626
QY 231 InProAlaGlnGlnAlaAlaLeuGlnLysGlnGlnValValMetAlaGluAlaAs 251
Db 627 AGCTTGCCACACAGCAGGAGCATTTACAGAAACAACAGGAAGTAGTAGTGGCTGGGTCTT 686
QY 251 exLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnG 271
Db 687 CCTTGCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAAGTTAATG 746
QY 271 lyGln-AlaLysThrHisThrGluAsnSerGluLysValLeuGlu 285
Db 747 GACAGGGCCAAAACACACACTGCACACTCCGAAAAAGAACTGGAA 791

RESULT 15
BU361227
LOCUS 603790024F1 CSEQCHN72 Gallus gallus cDNA clone CHEST75205 5', mRNA
DEFINITION sequence.
ACCESSION BU361227
VERSION BU361227.1 GI:25869228
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1. (bases 1 to 773)
AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
Location/Qualifiers
1. .773
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton line 151"
/db_xref="taxon:9031"
/clone="CHEST75205"
/sex="Female"
/tissue_type="cerebrum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN72"
/notes="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methyated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 18:56:48 ; Search time 244 Seconds  
(without alignments)  
3104.902 Million cell updates/sec

Title: US-09-762-594-7

Perfect score: 2394

Sequence: 1 RRRPKVELFTVTRVKVVL.....DNSYSLMRKSVYRYVYTR 463

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09762594/runat\_19082005\_171346\_18229/app\_query.fasta\_1.647  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09762594 @CEN 1.1.69 @runat\_19082005\_171346\_18229 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	392.5	16.4	432	4	US-09-270-767-26802
C 2	392.5	16.4	569	4	US-09-270-767-11257
C 3	229	9.6	3798	4	US-09-688-1888-9
C 4	229	9.6	3798	4	US-09-291-417D-9
5	226.5	9.5	4266	3	US-09-651-011A-3
6	209.5	8.8	1835	4	US-09-216-393B-80
7	207.5	8.7	4133	4	US-09-688-1888-11
8	207.5	8.7	4133	4	US-09-291-417D-11
9	206	8.6	1257	4	US-09-248-796A-7812
10	198.5	8.3	1952	4	US-09-520-781-23
11	198.5	8.3	2010	4	US-09-520-781-21
12	194.5	8.1	1542	4	US-09-248-796A-4611

13	193.5	8.1	3536	4	US-09-949-016-1083	Sequence 1083, Ap
14	193	8.1	3894	4	US-09-645-456A-7	Sequence 7, Appli
15	193	8.1	3894	4	US-09-425-324A-7	Sequence 7, Appli
16	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
17	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
18	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
19	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
20	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
21	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
22	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
23	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
24	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
25	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
26	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
27	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
28	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
29	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
30	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
31	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
32	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
33	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
34	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
35	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
36	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
37	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
38	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
39	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
40	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
41	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
42	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
43	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
44	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli
45	193	8.1	3894	4	US-09-645-791-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-270-767-26802/c  
; Sequence 26802, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 26802  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-26802

Alignment Scores:

Pred. No.: 2,09e-34 Length: 432  
Score: 392.50 Matches: 76  
Percent Similarity: 66.43% Conservative: 17  
Best Local Similarity: 54.29% Mismatches: 31  
Query Match: 16.40% Indels: 16  
DB: 4 Gaps: 2

US-09-762-594-7 (1-463) x US-09-270-767-26802 (1-432)

Qy	340	ValArgValProThrHisGluGlu-GlySerTyrLeuPheTrpGluPheAlaThrAspSe	359
Db	430	GTCCGTGTGCCCCAACCAACATGACGGCAAGTCATCTCTGGAGTTCGCCACGACAC	371
Qy	359	rTyrAspIleGlyPheGlyValTyrPheGluTrpThrAspSerProAsnAlaValSe	379
Db	370	CTACGACATCGGCTTCGGCATTTACTTCGAGTGGGCGCAAGCCCGTACCAACGAGGTGAC	311





Qy 268 rValAsnGlyCln---AlaLysThrHisThrGluAsnSerGlu----- 281  
 Db 1710 GGTGAAGATAGATTAGGAAATAACACAGCTCCCTGAGCCAGCTTAAGCAGAC 1769  
 Qy 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLy 298  
 Db 1770 AGGCAGAGTATTGGAGCACCAGTGCCTCCCGATCAGAGTCTTTTCCATGCAACTC 1829  
 Qy 298 sAspSerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317  
 Db 1830 CGAGTCTGTGATCCCGCCCTGCAGACACAGCG-----GAGCCACAGGTT----- 1875  
 Qy 317 pPheLysGluLysIleArgGlnAspAlaSerValIleThrValArgArg 334  
 Db 1876 -----CCTGTGAAGAACATCTCGCTCCCTGTTCTGTGTCGTCGA 1917

## RESULT 4

US-09-291-417D-9

; Sequence 9, Application US/09291417D

; Patent No. 6680170

; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY

; APPLICANT: MARTINEZ, RICARDO

; APPLICANT: WYTHE, DAVID

; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES

; FILE REFERENCES: 038602/0329

; CURRENT APPLICATION NUMBER: US/09/291.417D

; CURRENT FILING DATE: 1999-04-13

; PRIOR APPLICATION NUMBER: 60/081,784

; PRIOR FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 155

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 9

; LENGTH: 3798

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-291-417D-9

## Alignment Scores:

Pred. No.: 8,98e-15 Length: 3798  
 Score: 229.00 Matches: 76  
 Percent Similarity: 48.45% Conservative: 49  
 Best Local Similarity: 29.46% Mismatches: 78  
 Query Match: 9.57% Indels: 56  
 DB: 4 Gaps: 10

US-09-762-594-7 (1-463) x US-09-291-417D-9 (1-3798)

Qy 115 ArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArgArgGlnArg 134  
 Db 1198 CGGATTGAGCAGCAGAAAGAACAGAGCGCGCTAGAACAGCAACAAGAGAGAGCGG 1257  
 Qy 135 Glu-----GluGluGluArgGluArgLysGluGluLysArgLys 150  
 Db 1258 GAAGCTAGAGGCGAGCAACGTGAACAGCGAAGGAGAGAACAAAGAAAGAGCGCT 1317  
 Qy 151 ArgGluGluLysArgLeuArgArgGluGluGluArgArgIleGluGluGlu 170  
 Db 1318 CTAGAGAGGTGGAGAGAGCGCCAAAGAAAGAGAGAGAGAGAGCGCGCAGAAAGAA 1377  
 Qy 171 ArgLeuArgLeuGluGlnLysGlnIleMetAlaLeuAsn----- 186  
 Db 1378 AAGAGGAGAGTTGAAGAGAACAGGAGTATATCAGCGCAGCAGCTAGAACAGGAGCGG 1437  
 Qy 187 -----SerGlnThrAlaValGlnPheGln----- 194  
 Db 1438 CACTTGGAGTCTCTCAGCAGCAGCTGCTCCAGGAGCAGGCCATGTTACTGGAGTGC 1497  
 Qy 195 -----GlnTyAlaAlaGlnGlnTyProGlyAsnTyArgGluGln 208  
 Db 1498 TGGCGGAGATGAGGAGCACCGCGCAGCAGAGGCTCCAGAGCGAGTTCGCAACAGAA 1557  
 Qy 209 GlnIleLeuIleArgGlnLeuGlnGlnHisTyArgGlnIleTyLysHisGlnAlaGlu 228

Db 1558 CAAGCATATCTCTGTCTCTACAGCATGACCATAGGAGCGCGACCGCAGCACTCGCAG 1617  
 Qy 229 GlnThrGlnProAlaGlnGlnGln-AlaAlaLeuGlnLysGlnGlnGluValValMetAl 248  
 Db 1618 CAGCGCCACACCGCAGCAGGAAAGAGCAAGCCAA----- 1654  
 Qy 248 aGlyAlaSerLeuProAlaSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268  
 Db 1655 -----GCTTCATGCTCCGAGCCCAAGCCCACTACGAGCT-GCTGACCGAGCGCGAGA 1709  
 Qy 268 rValAsnGlyCln---AlaLysThrHisThrGluAsnSerGlu----- 281  
 Db 1710 GGTGAAGATAGATTAGGAAATAACACAGCTCCCTGAGCCAGCTTAAGCAGAC 1769  
 Qy 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLy 298  
 Db 1770 AGGCAGAGTATTGGAGCACCAGTGCCTCCCGATCAGAGTCTTTTCCATGCAACTC 1829  
 Qy 298 sAspSerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317  
 Db 1830 CGAGTCTGTGATCCCGCCCTGCAGACACAGCG-----GAGCCACAGGTT----- 1875  
 Qy 317 pPheLysGluLysIleArgGlnAspAlaSerValIleThrValArgArg 334  
 Db 1876 -----CCTGTGAAGAACATCTCGCTCCCTGTTCTGTGTCGTCGA 1917

## RESULT 5

US-09-651-011A-3

; Sequence 3, Application US/09651011A

; Patent No. 6346416

; GENERAL INFORMATION:

; APPLICANT: Nicholas M. Dean

; APPLICANT: Lex M. Cowsett

; TITLE OF INVENTION: ANTISENSE MODULATION OF HPK/GCK-LIKE KINASE EXPRESSION

; FILE REFERENCE: RTS-0168

; CURRENT APPLICATION NUMBER: US/09/651,011A

; CURRENT FILING DATE: 2000-08-29

; NUMBER OF SEQ ID NOS: 49

; SEQ ID NO 3

; LENGTH: 4266

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(3528)

US-09-651-011A-3

## Alignment Scores:

Pred. No.: 2,04e-14 Length: 4266  
 Score: 226.50 Matches: 71  
 Percent Similarity: 49.58% Conservative: 46  
 Best Local Similarity: 30.08% Mismatches: 76  
 Query Match: 9.46% Indels: 43  
 DB: 3 Gaps: 10

US-09-762-594-7 (1-463) x US-09-651-011A-3 (1-4266)

Qy 115 ArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArgArgGlnArg 134  
 Db 1081 CGGATTGAGCAGCAGAAAGAACAGAGCGCGCTAGAACAGCAACAAGAGAGAGCGG 1140  
 Qy 135 Glu-----GluGluGluArgGluArgLysGluGluLysArgLys 150  
 Db 1141 GAAGCTAGAGGCGAGCAACGTGAACAGCGAAGGAGAGAACAAAGAAAGAGCGCT 1200  
 Qy 151 ArgGluGluLysArgLeuArgArgGluGluGluArgArgIleGluGluGlu 170  
 Db 1201 CTAGAGAGTGGAGAGAGCGCCAAAGAAAGAGAGAGAGAGAGCGCGCAGAAAGAA 1260  
 Qy 171 ArgLeuArgLeuGlnGlnLysGlnIleMetAlaLeuAsnSerGlnThrAla 190  
 Db 1261 AAGAGGAGAGTTCGAAAGAGAA----- 1281



Percent Similarity: 48.97% Conservative: 51  
Best Local Similarity: 27.98% Mismatches: 93  
Query Match: 8.67% Indels: 31  
DB: 4 Gaps: 8

US-09-762-594-7 (1-463) x US-09-688-188B-11 (1-4133)

QY 78 ArgArgGluTrpAlaAlaLeu-----GlyAsnMetSerLysGluAspAlaMetValGlu 95  
Db 952 CGCGGGAGTTTCTCCGGCTCCAGCAGGAAATAAGACCAACTCAGAGGCTTTAAACAG 1011  
QY 96 PheValLysLeuAsnLysCysProLeuLeuSerAlaTyrVal-----111  
Db 1012 CAGCAGAGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1071  
QY 112 -----AlaSerHisArgGluGluGluGluGluGluGluGluGluGluGluGlu 129  
Db 1072 CACCAGCGCGCGCATAGAGGAGCAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCG 1131  
QY 130 GluArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 149  
Db 1132 CAGCGCGGAGCGCGGAGCAG-----CGGAAGCTGCAGGAGGAGCAGCAGCGG 1182  
QY 150 LysArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 165  
Db 1183 CGCGTGGAGGACATCAGCGCTCTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1242  
QY 166 -----ArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 179  
Db 1243 CAGGAATATATTCGTACAGGCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1302  
QY 180 GlnIleMetAlaAlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyrAlaAlaGln 199  
Db 1303 CAGCTGCTC-----CAGGAACAGCGCTCTGCTGGAATACAAAGCGGAAG 1347  
QY 200 GlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeuIleArgGlnLeuGlnGlnHis 219  
Db 1348 CAGCTGGAG---GAGCAGCGGCGAGTCAGAACGCTCTCCAGAGGCGAGCTGCAGCAGGAGCAT 1404  
QY 220 TyrGlnGlnTyrLysHisGlnAlaGluGlnThrGlnProAlaGlnGlnGlnAlaLeu 239  
Db 1405 -----GCCTACTCAAGTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1452  
QY 240 GlnLysGlnGlnGluValValMetAlaGlyAlaSerLeuProAlaSerLysValAsn 259  
Db 1453 AAACAACAGCAGCAGCAGCTCTCGCTGGGAGCAGGAGCCCTGTACCATTTATGTCGG 1512  
QY 260 ThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsn 279  
Db 1513 GGCATGAATCCCGCTGACAAACCGAGCTGGCGCCGAGAGGTAGAGAGAGAGAGAGAGATG 1572  
QY 280 SerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyProLysAsp 299  
Db 1573 AACAGCAGCAGCAACTCTCTCCCTTGGCCAAAGACGAGCAGCAGCAGCGCGGCTGAGCC 1632  
QY 300 SerLeuPro 302  
Db 1633 CCCATCCCC 1641

## RESULT 8

US-09-291-417D-11  
; Sequence 11, Application US/09291417D  
; Patent No. 6680170  
; GENERAL INFORMATION:  
; APPLICANT: FLOWMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHYTE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 038602/0329  
; CURRENT APPLICATION NUMBER: US/09/291,417D  
; CURRENT FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: 60/081,784  
; PRIOR FILING DATE: 1998-04-14

## RESULT 9

US-09-248-796A-7812  
; Sequence 7812, Application US/09248796A

; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: PatentIn Ver. 2.11  
; SEQ ID NO 11  
; LENGTH: 4133  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-291-417D-11

Alignment Scores:  
Pred. No.: 2,51e-12 Length: 4133  
Score: 207.50 Matches: 68  
Percent Similarity: 48.97% Conservative: 51  
Best Local Similarity: 27.98% Mismatches: 93  
Query Match: 8.67% Indels: 31  
DB: 4 Gaps: 8

US-09-762-594-7 (1-463) x US-09-291-417D-11 (1-4133)

QY 78 ArgArgGluTrpAlaAlaLeu-----GlyAsnMetSerLysGluAspAlaMetValGlu 95  
Db 952 CGCGGGAGTTTCTCCGGCTCCAGCAGGAAATAAGACCAACTCAGAGGCTTTAAACAG 1011  
QY 96 PheValLysLeuAsnLysCysProLeuLeuSerAlaTyrVal-----111  
Db 1012 CAGCAGCAGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1071  
QY 112 -----AlaSerHisArgGluGluGluGluGluGluGluGluGluGluGluGlu 129  
Db 1072 CACCAGCGCGCGCATAGAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1131  
QY 130 GluArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 149  
Db 1132 CAGCGCGGAGCGCGGAGCAG-----CGGAAGCTGCAGGAGGAGCAGCAGCGG 1182  
QY 150 LysArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 165  
Db 1183 CGCGTGGAGGACATCAGCGCTCTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1242  
QY 166 -----ArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 179  
Db 1243 CAGGAATATATTCGTACAGGCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1302  
QY 180 GlnIleMetAlaAlaLeuAsnSerGlnThrAlaValGlnPheGlnGlnTyrAlaAlaGln 199  
Db 1303 CAGCTGCTC-----CAGGAACAGCGCTCTGCTGGAATACAAAGCGGAAG 1347  
QY 200 GlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeuIleArgGlnLeuGlnGlnHis 219  
Db 1348 CAGCTGGAG---GAGCAGCGGCGAGTCAGAACGCTCTCCAGAGGCGAGCTGCAGCAGGAGCAT 1404  
QY 220 TyrGlnGlnTyrLysHisGlnAlaGluGlnThrGlnProAlaGlnGlnGlnAlaLeu 239  
Db 1405 -----GCCTACTCAAGTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1452  
QY 240 GlnLysGlnGlnGluValValMetAlaGlyAlaSerLeuProAlaSerLysValAsn 259  
Db 1453 AAACAACAGCAGCAGCAGCTCTCGCTGGGAGCAGGAGCCCTGTACCATTTATGTCGG 1512  
QY 260 ThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsn 279  
Db 1513 GGCATGAATCCCGCTGACAAACCGAGCTGGCGCCGAGAGGTAGAGAGAGAGAGAGAGATG 1572  
QY 280 SerGluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyProLysAsp 299  
Db 1573 AACAGCAGCAGCAACTCTCTCCCTTGGCCAAAGACGAGCAGCAGCAGCGCGGCTGAGCC 1632  
QY 300 SerLeuPro 302  
Db 1633 CCCATCCCC 1641



Qy	173	ArgLeuGluGlnGlnLysGlnGlnIleMetAlaAlaLeuAenSerGlnThrAlaValGln	192
Db	816	GAGAGAGAGACAAAGAACCGGAACGCCCTGGAGCGGCAGAACGCCTGGAG	875
Qy	193	PheGlnGlnTyrrAlaAlaGlnGlnTyrrProGlyAenTyrrGluGlnGlnIleLeulle	212
Db	876	CGG CAGGNA-----CGCCTGGAGCCGCAAGAACGCCTGGAT	911
Qy	213	ArgGlnLeuGlnGlu---GlnHisTyrrGlnGlnTyrrLysHisGlnAlaGluGlnThrGln	231
Db	912	CGGAGAGGCAAGAAAGACAAGACAGAGAGAGCGGTGGAGACACTGGAAACGGGAGAGCAA	971
Qy	232	ProAlaGlnGlnGlnAlaAlaLeuGlnLysGlnGln-----GluValVal	246
Db	972	GAAGGGAGCCCAAGAGACAGTATTGAAGAGGGAACAGCTGGNAATGGGAGAGAGCGCAGA	1031
Qy	247	MetaAlaGlyAlaSerLeuProAlaSerSerLysValAenThrAlaGlyAlaSerAspThr	266
Db	1032	ATATCAAGTGCTGTGCCCCCTGCCTCTGTTGAG-----ACTCCTCTA	1073
Qy	267	LeuSerValIaenGlyGlnAlaLysThrHisThrGluAenSerGluLysValLeuGluPro	286
Db	1074	AAC TCTGTGTGGGAGACTTCTTCGCTTGAGCCAGCGCTTGGAGGAGCGCTCTCACCGC	1133
Qy	287	--GluAlaAlaGluGluAlaLeuGluAenGlyProLysAepSerLeuProValIleAla	305
Db	1134	GCGGAGACTCCATCCAAACAGGAGAGACANTGCCCTTTAACTGACATTGCAGCTGCATTT	1193
Qy	306	Ala 306	
Db	1194	GCC 1196	

QY	113	SerHisArgIleGluLysGluGluGluGluLysArgGlyAsnThrAlaValGlnThrLeu	132
DB	639	AATGGCCCATCCCAAGAAGAAATTGGAATAATCAAAGAAGACAACACTACAAGAACACAGCAACGG	698
QY	133	GlnArgGluGluGluGluAurGluAurGluGluGluGluGluGluLysArgLysAlaValGlnThrLeu	152
DB	699	CAAAAG---GAGCTGGAGCGGGAAGGCTGAAGCGAGAAAAGAAATGGAAAGAGAAAGGAAG	755
QY	153	GluGluAspArgLeuAurGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	172
DB	756	AAGAGAGAGAGGTTAGAAGAGGAAGAGCTGGAGAGGAGCAGCTGGAAACCAAGAACACAGCTG	815
QY	173	ArgLeuGluGlnGlnLysGlnInileMetAlaAlaLeuAsnSerGlnThrAlaValGln	192
DB	816	GAGAGAGAGAGACAAGAAGAACGGAACGCCTGGAGCGGCAGGAACGCCCTGGAG	875
QY	193	PheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeulle	212
DB	876	CGGCAGGAA-----CGCCTGGAGCGGCAGGAACGCCCTGGAT	911
QY	213	ArgGlnLeuGlnGlu---GlnHisTyrGlnGlnTyrLyshiGlnAlaGluGlnThrGln	231
DB	912	CGGAGAGGACGAAGAACAAGACAGAGAGAGCTGGAGAGACTGGNACCGGAGAGGCAA	971
QY	232	ProAlaGlnGlnAlaAlaLeuGlnLysGlnGln-----GluValVal	246
DB	972	GAAGGGGAGCGCACAGAGCAGCTTAGAAGGGAACAGCTGGAATGGGAGAGAGCGCAGA	1031
QY	247	MetalGlyAlaSerLeuProAlaSerSerLysValasnThrAlaGlyAlaSerAspThr	266
DB	1032	ATATCAAGTGCTGTCGCCCTGCTCTGTTGAG-----ACTCCTCTA	1073
QY	267	LeuSerValaenGlyGlnAlaLysThrHisThruasnSerGluLysValLeuGluPro	286
DB	1074	AACCTCTGTCTGGAGACTCTCTGCTTCTGAGCCAGGCTTGCAGGAGCGCTCTCAGCCG	1133
QY	287	---GluAlaAlaGluGluAlaLeuGluAsnGlyProLysAspSerLeuProValIleAla	305
DB	1134	GCGGAGACTCCATCCCAACAGGAACAATCGCCCTTAACCTGACTTGCAGCTGCGCAATT	1193
QY	306	Ala 306	
DB	1194	GCC 1196	
<b>RESULT 12</b>			
US-09-248-796A-4611			
; Sequence 4611, Application US/09248796A			
; Patent No. 6747137			
; GENERAL INFORMATION:			
; APPLICANT: Keith Weinstock et al			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDID			
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.132			
; CURRENT APPLICATION NUMBER: US/09/248,796A			
; CURRENT FILING DATE: 1999-02-12			
; PRIOR APPLICATION NUMBER: US 60/074,725			
; PRIOR FILING DATE: 1998-02-13			
; PRIOR APPLICATION NUMBER: US 60/096,409			
; PRIOR FILING DATE: 1998-08-13			
; NUMBER OF SEQ ID NOS: 28208			
; SEQ ID NO 4611			
; LENGTH: 1542			
; TYPE: DNA			
; ORGANISM: Candida albicans			
US-09-248-796A-4611			
<b>Alignment Scores:</b>			
Pred. No.: 1,53e-11 Length: 1542			
Score: 194.50 Matches: 84			
Percent Similarity: 40.28% Conservative: 61			
Best Local Similarity: 23.33% Mismatches: 100			
Query Match: 8.12% Indels: 115			
DB: 4 Gaps: 17			



Qy	243	GlnGluValValMetAlaGlyAlaSerLeuProAla-----SerSerLysValAsnThr	260
		:::	
Db	2371	CTGAATGTGACTGGGACGTGCAGCTCT---CCAGCTTTGTACTCTCATCTCCATCCTCCG	2427
		:::	
Qy	261	AlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaIysThrH1sThrGluAsnSer	280
		:::	
Db	2428	CAAGGGCACAAAGGCCCTCTCCCAAGATCAACCCCACTACGGGATGGATCTCGAATAGC	2487
		:::	
Qy	281	GluLysValLeuGluProGluAlaAlaGluGluAlaLeuGluAsnGlyProLysAspSer	300
		:::	
Db	2488	GACGACTCCACCGATGATGAGGCC-----CATCCCGGAGAGCCC	2526
		:::	
Qy	301	LeuProValIleAlaAlaProSerMetTrpThrArg	312
		:::	
Db	2527	ATCCCCACCC-----TGGGCCCGGA	2544
		:::	

```

RESULT 14
US-09-645-456A-7
; Sequence 7, Application US/09645456A
; Patent No. 6562580
; GENERAL INFORMATION:
; APPLICANT: Fu, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: A-68344/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/645,456A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US/09/425,324
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 3894
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-645-456A-7

```

Alignment Scores:	
Pred. No.:	9.38e-11
Score:	193.00
Percent Similarity:	49.57%
Best Local Similarity:	27.53%
Query Match:	8.06%
DB:	4
	Gaps:
	Indels:
	Mismatches:
	Conservative:
	Matches:
	Length:

US-09-762-594-7 (1-463) x US-09-645-456A-7 (1-3894)

Qy	115	ArgIleGluLysGluGluGluGluLeuLysAerArgLysAlaGluGluGluAerArgGlnArg	134
Db	1180	CGCATCGAGCAGCAAAAGACGAGCGCGCTGGAGGAGCAACAAGCGCGAGAGAAG	1239
Qy	135	Glu-----GluGluGluAerGluAerGluLeuClnLysGluGluGluLysAerGlyAerG	151
Db	1240	GAGCTCGGAAGCAGCAGAGAGGAGCAGCGCGGCACCTATGAGGAGCAGATGCGCCGG	1299
Qy	152	GluGluAspArgLeuAerGluGluGluGluGlu-----ArgAerArgIleGluGlu	169
Db	1300	GAGGAGAGAGGAGGCGTGTGGGCACATGAACAGAGATACATCAGCGCACAGTTAGAGGAG	1359
Qy	170	GluAerGluAerGluGluGlnGlnLysGlnGlnIleMet-----AlaAlaLeuAsn	186
Db	1360	GAGCAGAGACAGTTTAGAGACTTTCGACGACGACGCTACTGCGATGAACAAGCTCTACTCTG	1419
Qy	187	SerGlnThrAlaValGlnPhe---GlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyr	205
Db	1420	GAATATAAGCGCAAACAATTGGGAAGAACACAGACAGCAGAAAGACTGTCAGAGCGACGCTA	1479
Qy	206	GluGlnGlnGlnIleLeuIleAerGlnLeuGlnGlnGluHisTyrGlnGln-----	222
Db	1480	AAGCAGAAAAGAGACTACTATTAGTTTCCTTTCCATCATCAGCGCGGAGCAGAGCGCCTGTG	1539

Qy	223	-----TyrLysHisGlnAlaGluGlnThrGlnProAlaGlnGlnAlaAa	238
Db	1540	GAGAAGAAGCCACTGTACCATTCACAAAGAAGTAAGTCTTAGTGAGAAGCCAGCATGG	1599
Qy	239	LeuGlnLysGlnGlnGluValValMetAlaGlyAlaSerLeuProAlaSerSerLysVal	258
Db	1600	GCCAAGGAGATCCACATCTGGTAGCTGTAAATATCCAGGACCTGCTTGCACGCCCTCC	1659
Qy	259	AsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGlu	278
Db	1660	CAGTCAGTCACGACGACGCCCAAAAGGGCTCTCTGGGTTTCAGGAGCTCTGAACGTG	1719
Qy	279	AsnSerGluLysValLeuGluPro--GluAlaAlaGluGluAlaLeuGluAsnGlyPro	297
Db	1720	ACCTCCACCGCTGGAGATGCCACGCCAACTCAGATCCACCTCGGAAATCCTCCT	1779
Qy	298	LysAspSerLeuProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAsp	317
Db	1780	CTCCCCACTCGCATTGAAAGATTGTGACGAAGCTCTTGGTTACGA--CAGGAAGAAGAC	1836
Qy	318	PhelysGluLysIleArgGlnAspAlaAspSerVal	329
Db	1837	ATTCACCAAGGTGCTCTCAAGAACAACTTCTATA	1872

```

RESULT 15
US-09-425-324A-7
; Sequence 7, Application US/09425324A
; Patent No. 6562591
; GENERAL INFORMATION:
; APPLICANT: FU, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER K
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: A-68344/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/425,324A
; CURRENT FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1

```

Alignment Scores:		
Pred. No.:	9.38e-11	Length: 3894
Score:	193.00	Matches: 64
Percent Similarity:	49.57%	Conservative: 51
Best Local Similarity:	27.53%	Mismatches: 99
Query Match:	8.06%	Indels: 18
DB:	4	Gaps: 7

US-09-762-594-7 (1-463) x US-09-425-324A-7 (1-3894)

Qy	115	ArgIleGluLysGluGluGluGluLysArgArgLysArgAlaGluGluGluArgArgGlnArg	134
Db	1180	CGCATCAGGACAGAAAGACAGAGCCGCCCTGGAGGAGCAACAAGAGCCGAGAGAAG	1239
Qy	135	Glu-----GluGluGluArgGluArgLeuGlnLysGluGluGluLysArgLysArg	151
Db	1240	GAGTCGGGAAGCAGCAGGAGAGGGAGCACGCCGCCACTATCAGGAGCAGATGCGCCGG	1299
Qy	152	GluGluGluAspArgLeuArgArgGluGluGluGlu-----ArgArgArgIleGluGlu	169
Db	1300	GAGGAGAGAGGAGGGCGTGGGAGCATGAACAGCNATACATCAGGCACAGTTAGAGGAG	1359
Qy	170	GluArgLeuArgLeuGlnGlnLysGlnGlnIleMet-----AlaAlaLeuAsn	186
Db	1360	GAGCAGAGACAGTTAGATGATCTTGACAGCAGCAGCTACTGCATGAACAAGCTTACTTCG	1419
Qy	187	SerGlnThrAlaValGlnPhe---GlnGlnTyAlaAlaGlnGlnTyProGlyAsnTyr	205



```
Db 1420 GAATATAAGCGCAACAATTGGAAGAACAGAGACAAGAGAACTGCAGAGGCAGCTA 1479
Qy 206 GluGlnGlnGlnLeuIleArgGlnLeuGlnGlnGlnHisTyrglnGln----- 222
Db 1480 AAGCAAGAAGAGACTACTAGTTTCCTTCAGCATCAGCGGAGAGAGAGCGCTGTG 1539
Qy 223 -----TyrglnHisGlnAlaGlnGlnThrGlnProAlaGlnGlnAlaA 238
Db 1540 GAGAAGAAGCCACTGTACCATTAACAAGAAGGATGAGTCCTAGTGAGAAGCCAGCATGG 1599
Qy 239 LeuGlnLysGlnGlnGluValMetAlaGlyAlaSerLeuProAlaSerSerLysVal 258
Db 1600 GCCAAGGAGATCCACATCTGGTAGCTGTAAATCCAGGGACCTGCTTGACCGCTCC 1659
Qy 259 AsnThrAlaGlyAlaSerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGlu 278
Db 1660 CAGTCAGTGCAGCAGCAGCCCAAAAGGGCCTCTCTGGGTTTCAGGAGGCTCTGAACGTG 1719
Qy 279 AsnSerGluLysValLeuGluPro---GluAlaAlaGluAlaLeuGluAsnGlyPro 297
Db 1720 ACCTCCACCGGTGGAGATGCCACGCCAGAACTCAGATCCACCTCGGAAATCCTCT 1779
Qy 298 LysAspSerLeuProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAsp 317
Db 1780 CTCCCACTCGCATTGAAAAGTTTGACCGAGAGCTCTTGGTTACGA---CAGGAAGAAGAC 1836
Qy 318 PheLysGluLysIleArgGlnAspAlaAspSerVal 329
Db 1837 ATTCCACCAAGGTGCCTCAAAAGACACAACTTCTATA 1872
```

Search completed: August 23, 2005, 00:08:12  
Job time : 257 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 08:25:19 ; Search time 269 Seconds  
(without alignments)  
8874.824 Million cell updates/sec

Title: US-09-762-594-2  
Perfect score: 1459  
Sequence: 1 gaattcgccgcgctgcac.....catgacgttaatttccttt 1459

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	99.6	6.8	7218	1	US-08-232-463-14
2	78.4	5.4	486	4	US-09-639-207-13
3	74.2	5.1	3489	2	US-08-728-323A-1
4	74.2	5.1	3489	3	US-09-298-568-1
5	74.2	5.1	3489	4	US-09-410-399-1
6	74.2	5.1	3489	4	US-09-894-273-1
C 7	74.2	5.1	32207	2	US-08-770-379-20
C 8	74.2	5.1	32207	3	US-08-757-669A-20
C 9	74.2	5.1	32207	3	US-09-230-371A-20
C 10	73.6	5.0	12695	4	US-09-949-016-16775
11	72.6	5.0	1926	3	US-09-249-585A-2
12	72.6	5.0	1926	4	US-09-410-399-3
13	72.6	5.0	2580	3	US-09-050-863-2
14	72.6	5.0	2580	3	US-09-359-081-2
C 15	72.6	5.0	5452	2	US-09-130-114-1
C 16	72.6	5.0	8705	4	US-09-647-344A-14
17	72.6	5.0	9600	3	US-08-910-647-1
18	72.6	5.0	9600	3	US-09-620-925-1
19	72.6	5.0	10596	1	US-07-884-811-15
20	72.6	5.0	10596	1	US-07-885-971-15
21	72.6	5.0	10596	1	US-08-087-783A-15
22	72.6	5.0	10596	1	US-08-194-088B-15
23	72.6	5.0	10596	2	US-08-194-087-15
24	72.6	5.0	10596	5	PCT-US93-04648-15
C 25	72.6	5.0	16080	4	US-09-724-566A-48
C 26	72.6	5.0	16080	4	US-09-471-669A-48
27	70.8	4.9	543	6	5273901-6

28	70.8	4.9	543	6	5273901-6
29	69.4	4.8	533	6	5482709-5
30	69.4	4.8	533	6	5482709-5
C 31	69.4	4.8	16442	3	US-08-781-891-208
C 32	69.4	4.8	16442	4	US-09-618-166-208
33	67.8	4.6	767677	4	US-09-949-016-12147
34	67.8	4.6	767677	4	US-09-949-016-17361
C 35	65.6	4.5	51259	3	US-08-781-891-209
C 36	65.6	4.5	51259	4	US-09-618-166-209
C 37	64.6	4.4	432	4	US-09-270-767-26802
C 38	64.6	4.4	569	4	US-09-270-767-11257
39	63.6	4.4	9551	1	US-08-056-200-93
40	63.6	4.4	9551	2	US-08-800-644-93
C 41	63	4.3	85850	4	US-09-949-016-13424
42	62.8	4.3	1758	4	US-09-949-016-1756
43	62.8	4.3	194937	4	US-09-949-016-17032
44	62.8	4.3	194937	4	US-09-949-016-17033
45	62.4	4.3	601	4	US-09-949-016-32546

ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ9pt-F1s  
; US-08-232-463-14  
Query Match 6.8%; Score 99.6; DB 1; Length 7218;



Db 1937 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 1996  
 QY 477 TGAGACGGAGGAGGAGAGAGAGCGCGGATAGAGGAAGAGAGGCTTCGGCTGGAACAGC 536  
 Db 1997 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2056  
 QY 537 AAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCAATTCACGACGATG 596  
 Db 2057 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2116  
 QY 597 CAGCCAGCAGTAT---CCAGGGAACCTACGAAACACAGCAGATTCCTATCCGCGACGTCG 653  
 Db 2117 AGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGC 2176  
 QY 654 AGCAGCAGCAGCTATCAGCAGTATAAACACACAG--GCAGAGCAAAACCACTGCACAACA 711  
 Db 2177 AGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG 2236  
 QY 712 ACAGGCAGCATTACAGAAACAGCAAGAGTAGTGCTGGGGGCATCATTTGCTGCATC 771  
 Db 2237 AGCAGCAGCAGCAGGATGAACAGCAGCAGCAGGAGCAGGAGCAGCAGGAGCAGGAGCAGG 2296  
 QY 772 ATCAAAGGTGAACACAGCTGGAGCAAGTAGTACACTGTCTCAGTTAATGGA 820  
 Db 2297 AGCAGGATGAGGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAG 2345

RESULT 4

US-09-298-568-1  
 ; Sequence 1, Application US/09298568  
 ; Patent No. 6322792  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kieff, Elliott D.  
 ; APPLICANT: Ballester, Mary E.  
 ; APPLICANT: Kaye, Kenneth M.  
 ; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
 ; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE  
 ; FILE REFERENCE: 16412-10001R  
 ; CURRENT APPLICATION NUMBER: US/09/298,568  
 ; CURRENT FILING DATE: 1999-04-21  
 ; EARLIER APPLICATION NUMBER: US 60/109,422  
 ; EARLIER FILING DATE: 1998-11-19  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3489  
 ; TYPE: DNA  
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
 US-09-298-568-1

Query Match 5.1%; Score 74.2; DB 3; Length 3489;  
 Best Local Similarity 50.3%; Pred. No. 1.7e-10;  
 Matches 236; Conservative 0; Mismatches 228; Indels 5; Gaps 2;  
 QY 357 AGAAGGAAG 416  
 Db 1877 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 1936  
 QY 417 AAGAAACGAGAGCGGCTGCAAAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476  
 Db 1937 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 1996  
 QY 477 TGAGACGGAGGAGGAG 536  
 Db 1997 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2056  
 QY 537 AAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCAATTCACGACGATG 596  
 Db 2057 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2116  
 QY 597 CAGCCAGCAGTAT---CCAGGGAACCTACGAAACACAGCAGATTCCTATCCGCGACGTCG 653  
 Db 2117 AGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGC 2176  
 QY 654 AGAAGCAGAGCGGCTGCAAAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476  
 Db 1937 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 1996  
 QY 477 TGAGACGGAGGAGGAG 536  
 Db 1997 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2056  
 QY 537 AAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCAATTCACGACGATG 596  
 Db 2057 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2116  
 QY 597 CAGCCAGCAGTAT---CCAGGGAACCTACGAAACACAGCAGATTCCTATCCGCGACGTCG 653

Db 2117 AGGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGC 2176  
 QY 654 AGCAGCAGCAGCTATCAGCAGTATAAACACACAG--GCAGAGCAAAACCACTGCACAACA 711  
 Db 2177 AGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG 2236  
 QY 712 ACAGGCAGCATTACAGAAACAGCAAGAGTAGTGCTGGGGGCATCATTTGCTGCATC 771  
 Db 2237 AGCAGCAGCAGCAGGATGAACAGCAGCAGCAGGAGCAGGAGCAGCAGGAGCAGGAGCAGG 2296  
 QY 772 ATCAAAGGTGAACACAGCTGGAGCAAGTAGTACACTGTCTCAGTTAATGGA 820  
 Db 2297 AGCAGGATGAGGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAG 2345

RESULT 5

US-09-410-399-1  
 ; Sequence 1, Application US/09410399  
 ; Patent No. 6482587  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robertson, Erle S.  
 ; APPLICANT: Cottler, Murray A.  
 ; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA  
 ; TITLE OF INVENTION: Methods to Genomic Host DNA  
 ; FILE REFERENCE: UM-03778  
 ; CURRENT APPLICATION NUMBER: US/09/410,399  
 ; CURRENT FILING DATE: 1999-10-01  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3489  
 ; TYPE: DNA  
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
 US-09-410-399-1

Query Match 5.1%; Score 74.2; DB 4; Length 3489;  
 Best Local Similarity 50.3%; Pred. No. 1.7e-10;  
 Matches 236; Conservative 0; Mismatches 228; Indels 5; Gaps 2;  
 QY 357 AGAAGGAAG 416  
 Db 1877 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 1936  
 QY 417 AAGAAACGAGAGCGGCTGCAAAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476  
 Db 1937 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 1996  
 QY 477 TGAGACGGAGGAGGAG 536  
 Db 1997 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2056  
 QY 537 AAAAGCAGCAGATTAATGGCAGCTTTAACTCGCAGACTGCCGTGCAATTCACGACGATG 596  
 Db 2057 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 2116  
 QY 597 CAGCCAGCAGTAT---CCAGGGAACCTACGAAACACAGCAGATTCCTATCCGCGACGTCG 653  
 Db 2117 AGGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGC 2176  
 QY 654 AGGAGCAGCAGCTATCAGCAGTATAAACACACAG--GCAGAGCAAAACCACTGCACAACA 711  
 Db 2177 AGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG 2236  
 QY 712 ACAGGCAGCATTACAGAAACAGCAAGAGTAGTGCTGGGGGCATCATTTGCTGCATC 771  
 Db 2237 AGCAGCAGCAGCAGGATGAACAGCAGCAGCAGGAGCAGGAGCAGCAGGAGCAGGAGCAGG 2296  
 QY 772 ATCAAAGGTGAACACAGCTGGAGCAAGTAGTACACTGTCTCAGTTAATGGA 820  
 Db 2297 AGCAGGATGAGGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAG 2345

RESULT 6













GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 22:38:38 ; Search time 882 Seconds  
(without alignments)  
3410.514 Million cell updates/sec

Title: US-09-762-594-7

Perfect score: 2394

Sequence: 1 RPRPKVELFTVTRVKVL.....DNSYSLWSKSVYRVYTR 463

Scoring table:

BLOSUM62 Xgapop 10.0, Xgapext 0.5  
Xgapop 10.0, Xgapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09762594/runat\_19082005\_171347\_18277/app\_query.fasta\_1.647  
-DB=Published Applications NA -QFMT=fastcap -SURFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09762594.CGN 1 1 480 @runat\_19082005\_171347\_18277  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10J\_PUBCOMB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2021	84.4	1481	21	US-10-998-342-9
2	509	21.3	882	18	US-10-363-616-238
3	506	21.1	3547	17	US-10-108-260A-832
4	384	16.0	278	19	US-10-469-289-327
5	352	14.7	934	9	US-09-866-285-33
6	309	12.9	210204	20	US-10-723-860-1746
7	304.5	12.7	554	9	US-09-864-761-13381
8	302	12.6	306	9	US-09-864-761-130285
9	230.5	9.6	4959	10	US-09-930-213-255
10	229	9.6	3792	20	US-10-820-583A-18
11	229	9.6	3798	10	US-09-291-417-9
12	229	9.6	3798	20	US-10-725-329-9
13	229	9.6	3798	22	US-10-725-121-9
14	228.5	9.5	3612	21	US-10-494-940-21
15	226.5	9.5	3925	16	US-10-247-671-58
16	226.5	9.5	3925	18	US-10-168-582-21
17	226.5	9.5	4197	22	US-10-934-998-284
18	226.5	9.5	4266	22	US-10-934-998-9
19	226.5	9.5	5816	20	US-10-723-860-7048
20	226.5	9.5	7171	20	US-10-820-583A-15
21	226.5	9.5	7495	20	US-10-820-583A-16
22	226.5	9.5	7986	19	US-10-618-941-66
23	216.5	9.0	2436	9	US-09-842-552-105
24	214	8.9	3817	16	US-10-269-909-48
25	214	8.9	7033	21	US-10-820-583A-17
26	214	8.9	7033	21	US-10-895-235A-4
27	212.5	8.9	4200	22	US-10-840-512-91
28	211	8.8	3822	10	US-09-789-390-10
29	210.5	8.8	1657	10	US-09-789-390-29
30	210.5	8.8	1725	10	US-09-789-390-36
31	210.5	8.8	1788	10	US-09-789-390-33
32	210	8.8	1657	18	US-10-425-114-9563
33	210	8.8	2406	18	US-10-424-599-53216
34	209.5	8.8	1835	9	US-09-216-393-80
35	209.5	8.8	1835	21	US-10-321-856-80
36	209.5	8.8	2991	21	US-10-741-849-6219
37	207.5	8.7	3864	16	US-10-021-660-61
38	207.5	8.7	3864	18	US-10-211-462-88
39	207.5	8.7	3864	22	US-10-466-162-39
40	207.5	8.7	3999	10	US-09-789-390-6
41	207.5	8.7	4133	10	US-09-291-417-11
42	207.5	8.7	4133	20	US-10-725-329-11
43	207.5	8.7	4133	22	US-10-725-121-11
44	204.5	8.5	1787	10	US-09-789-390-66
45	204	8.5	628	16	US-10-029-386-22859

# ALIGNMENTS

## RESULT 1

US-10-998-342-9  
; Sequence 9, Application US/10998342  
; Publication No. US20050106672A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: LAL, Preeti  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: GORGONE, Gina A.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: PATTERSON, Chandra  
; TITLE OF INVENTION: NEUROTRANSMISSION ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0551 PCT

```

; CURRENT APPLICATION NUMBER: US/10/998,342
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: US/09/720,530
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/091,667
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte Clone No: 998868
US-10-998-342-9

Alignment Scores:
Pred. No.: 1,21e-198 Length: 1481
Score: 2021.00 Matches: 385
Percent Similarity: 94.21% Conservative: 22
Best Local Similarity: 89.12% Mismatches: 25
Query Match: 84.42% Indels: 0
DB: 21 Gaps: 0

US-09-762-594-7 (1-463) x US-10-998-342-9 (1-1481)

QY 32 LysAspGlyLysAlaPheHisProThrTyrGluGluLysLeuLysPheValAlaLeuHis 51
DB 74 AAAGATGGCAAGCACTTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCAT 133
QY 52 LysGlnValLeuLeuGlyProTyrAsnProAspThrSerProGluValGlyPhePheAsp 71
DB 134 AAGCAAGTCTTATATGGGCCCATATAATCCAGACACTTGTCTGAGGTTGGATTCTTTGAT 193
QY 72 ValLeuGlyAsnAspArgArgGluTyrAlaAlaLeuGlyAsnMetSerLysGluAsp 91
DB 194 GTGTTGGGAGNATGACAGGAGGAGAGATGGCCAGCCCTGGGAAACAATGCTTAAGAGGAT 253
QY 92 AlaMetValGluPheValLysLeuLeuAsnLysCysCysProLeuLeuSerAlaTyrVal 111
DB 254 GCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTCATCTCTTTTCAACATATGTT 313
QY 112 AlaSerHisArgIleGluLysGluGluGluLysArgArgLysAlaGluGluArg 131
DB 314 GCGTCCCAAAAATAGAGAAGGAAGCAAGCAAGCAAAAAGCAGCAGGAGGAGGAGCA 373
QY 132 ArgGlnArgGluGluGluArgGluArgGluGlnLysGluGluLysArgLysArg 151
DB 374 AGCGCGGTGNAGAGGAGAGAGAGAACTCTGCCAAAGAGGAGGAGAGAAACGTAGGAGA 433
QY 152 GluGluGluAspArgLeuArgArgGluGluGluArgArgArgIleGluGluGluArg 171
DB 434 GAAGAAGAGGAAAGGCTTCGACGGCGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 493
QY 172 LeuArgLeuGluGlnGlnLysGlnIleMetAlaAlaLeuAsnSerGlnThrAlaVal 191
DB 494 CTTCCGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCAGAGCTGCCGTG 553
QY 192 GlnPheGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnGlnIleLeu 211
DB 554 CAGTTCCAGCAGTATGACGCCCAACAGATATCCAGGGAATACGACAGCAGCAATTCCTC 613
QY 212 IleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
DB 614 ATCCGCGAGTTGCAGGAGCAACACTATCAGCAGTATCAGCAGCAGCTGTATCAAGTCCAG 673
QY 232 ProAlaGlnGlnAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 251
DB 674 CTTGCACAGCAACAGCAGCATTACAGAAACAAACAGAGAGTAGTAGTGGCTGGCTTCC 733
QY 252 LeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerValAsnGly 271

```

```

DB 734 TTGCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTGAGTTAATGGA 793
QY 272 GlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAlaGluGlu 291
DB 794 CAGGCCAAACACACACTGACAGCTCCGAAAGAAAGAACTGGAACCAAGAGCTGCAGAGAA 853
QY 292 AlaLeuGluAsnGlyProLysAspSerLeuProValIleAlaAlaProSerMetTyrThr 311
DB 854 GCCCTGGAGAAATGGACCAAAAGAAATCTCTTCAGTAATAGCAGCTCCATCCATGTGGACA 913
QY 312 ArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSerValIleThr 331
DB 914 CGACCTCAGATCAAGAGACTTCAAGAGAGAGATTACAGCAGATGCAGATTCCTGTATTACA 973
QY 332 ValArgArgGlyGluValValThrValArgValProThrHisGluGluGlySerTyrLeu 351
DB 974 GTGGGCCGAGGAGAGTGGTCACTGTTCGAGTACCCACCATGAAGAGGATCATATCTC 1033
QY 352 PheTyrGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPheGluTyrThr 371
DB 1034 TTTTGGGAATTTGCCACAGACAATATGACATTTGGGTTTGGGGTGTATTTTGAATGGACA 1093
QY 372 AspSerProAsnAlaAlaValSerValHisValSerGluSerSerAspGluGluGlu 391
DB 1094 GACTCTCCAAACACTGCTGTACGCGTCATGTTCAGTGTCCAGCATGACGACGAGGAG 1153
QY 392 GluGluGluAsnValThrCysGluGluLysAlaLysLysLysLysLysLysLysLysLys 411
DB 1154 GAAGAAGAAACATCGGTTGTGAAGAGAAAGCCAAAGAAATGCCAAACAAGCCTTTGCTG 1213
QY 412 AspGluIleValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHis 431
DB 1214 GATGAGATTGTCCCTGTGTACCGGAGACTGTCTATGAGAGGAGGTGTATCTGCGAGCCAT 1273
QY 432 GlnTyrProGlyArgGlyValTyrLeuLysPheAspAsnSerTyrSerLeuTyrArg 451
DB 1274 CAATATCCAGGAGAGGAGTCTATCTCTCAAGTTTGACAACTCCTACTCTTTGTGGCGG 1333
QY 452 SerLysSerValTyrTyrArgValTyrTyrThrArg 463
DB 1334 TCAAAATCAGTCTACTACAGAGCTATTATATACAGA 1369

RESULT 2
US-10-363-616-238
; Sequence 238, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 238
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)..(882)
US-10-363-616-238

Alignment Scores:
Pred. No.: 1,38e-42 Length: 882
Score: 509.00 Matches: 115
Percent Similarity: 55.78% Conservative: 54
Best Local Similarity: 37.95% Mismatches: 84
Query Match: 21.26% Indels: 50
DB: 18 Gaps: 8

US-09-762-594-7 (1-463) x US-10-363-616-238 (1-882)

```

```

Qy 172 LeuArgLeuGluGlnGlnLysGlnGlnIleMetAlaLeuAenSerGlnThrAlaVal 191
Db 82 CTGCTCCTCACCCACAGAGCCACAGATGGTATCTCAGTGAAGTATGAGGATGCCACGAA 141
Qy 192 GlnPheGlnGlnTyrAlaAlaGlnGlnTyrProGlyAsnTyrGluGlnGlnIleLeu 211
Db 142 GATCTCGGGAAGCAACT-----GGTCTTTGGAGGCTCAGGCTTGGTG 186
Qy 212 IleArgGlnLeuGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGluGlnThrGln 231
Db 187 AAACAGGATTGCTG-----CCTGACAGCCAGGCCAG 219
Qy 232 ProAlaGlnGlnGlnAlaLeuGlnLysGlnGln-----GluValValMetAla 248
Db 220 GTCTCAATGATGAGTGAAGTATCAAGTATCCACAGAGGCTGCGGACATCGTTATGATC 279
Qy 249 GlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSer 268
Db 280 -----CAGTCTGAACATACAGAGGCTATAGATGTTCTTTCA 315
Qy 269 ValAsnGlnAlaLysThrHisThrGluAsnSerGluLysValLeuGluProGluAla 288
Db 316 GCTGAT-----TTGGAA----- 327
Qy 289 AlaGluGlnAlaLeuGluAenGlyProLysAspSerLeuProValIleAlaAlaProSer 308
Db 328 TCTGCAGATCTTCTGGGGACCCACAGAAAGTCTCCACCTCTGTATGGCTCTCCATGC 387
Qy 309 MetTrpThrArgProGlnIleLysAspPheLysGluLysIleArgGlnAspAlaAspSer 328
Db 388 ATCTGGACCTTTGCCAGGTTGAAGAAATTCANAAGCAAGCTGGGCAAGAGAAGACAGC 447
Qy 329 ValIleThrValArgArgGlyGluValValThrValArgValProThrHisGluGluGly 348
Db 448 CGTCTGCTGTGAGCGTGTGAGTGTGACCAATCCGGTACTCTACTCTATCCAGAGGG 507
Qy 349 SerTyrLeuPheTrpGluPheAlaThrAspSerTyrAspIleGlyPheGlyValTyrPhe 368
Db 508 AAGCGTCTCTCTGGAGATTTCGACCGATGACTATGACATTTGGCTTTGGAGTTATTTT 567
Qy 369 GluTrpThrAspSerProAsnAlaAlaValSerValHisValSerGluSerSer---Asp 387
Db 568 GACTGGACCCCTGTAACTAGCATGACATAACTGTGCAGGTGAGTTCAGTTCAGTACGAC 627
Qy 388 GluGluGluGluGluGluAenValThrCysGlu----- 399
Db 628 GAGATGAAGAGAGAGAGAGAGAGAGATTGAAGAACCCGTTCCAGCTGGAGATGTG 687
Qy 400 GluLysAlaLysLysAenAlaAenLysProLeuLeuAspGluIleValProValTyrArg 419
Db 688 GAGAGAGCTCCAGGAGCTCTTTCGGGGTTCGCTATGGGAGGTCTATGCTGTGTACCG 747
Qy 420 ArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrProGlyArgGlyValTyr 439
Db 748 CGGGACAGCCACGACGTCAGGCTGCGGCTGGCAGCATGACTACCTGCTGGTGGGCACTAC 807
Qy 440 LeuLeuLysPheAspAenSerTyrSerLeuTrpArgSerLysSerValTyrTyrArgVal 459
Db 808 CTGCTCAAGTTCGACAACTCTACTCTCTGCTGGCAACAAGACTCTCTACTTCCACATC 867
Qy 460 TyrTyrThr 462
Db 868 TACTACAC 876

```

## RESULT 3

```

US-10-108-260A-832
; Sequence 832, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106

```

```

; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 832
; LENGTH: 3547
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-832

```

```

Alignment Scores:      1,91e-41      Length:      3547
Score: No.:           506.00          Matches:     107
Percent Similarity:    60.48%         Conservative: 43
Best Local Similarity: 43.15%         Mismatches:   62
Query Match:          21.14%         Indels:       36
DB:                   17              Gaps:         6

```

US-09-762-594-7 (1-463) x US-10-108-260A-832 (1-3547)

```

Qy 227 AlaGluGlnThrGlnProAlaGlnGlnAlaAlaLeuGlnLysGlnGln----- 243
Db 345 GCAGACACGAGCCCGAGTCTCAATGATGAGTGGCTAAGTATCAAGTTCACAGAGGCTCTGGG 404
Qy 244 GluValValMetAlaGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAla 263
Db 405 GACATCGTTATGATC-----CAGTCTGAACATACAGGAGCT 440
Qy 264 SerAspThrLeuSerValAsnGlyGlnAlaLysThrHisThrGluAsnSerGluLysVal 283
Db 441 ATAGATGTTCTTTTCACTGAT----- 461
Qy 284 LeuGluProGluAlaAlaGluGluAlaLeuGluAenGlyProLysAspSerLeuProVal 303
Db 462 TTGGAA-----TCTGCAGATCTTCTGGGGACACAGGAAAGTCTCCACCTCTG 512
Qy 304 IleAlaAlaProSerMetTrpThrArgProGlnIleLysAspPheLysGluLysIleArg 323
Db 513 ATGGCTCTCTCATGCTCTGGACCTTTTGCAGAGTGAAGAAATTCAAAAGCAAGCTGGGC 572
Qy 324 GlnAspAlaAspSerValIleThrValArgArgGlyGluValValThrValArgValPro 343
Db 573 AAAGAGAAGAAGACAGCCGCTCTGGTGTGAAGCGTGTGAGGTGGTGACCATCCGGGTACCT 632
Qy 344 ThrHisGluGluGlySerTyrLeuPheAlaThrAspSerTyrAspIleGly 363
Db 633 ACTCATCCAGAGGGGAAGCGTGTCTCTGGAGTTTGGCACCAGTACTATGATCATTTGGC 692
Qy 364 PheGlyValTyrPheGluTrpThrAspSerProAsnAlaAlaValSerValHisValSer 383
Db 693 TTTGGAGTTTATTTTGACTGGACCCCTGTAACTAGCACTGACATACTGTGCAGGTCACT 752
Qy 384 GluSerSer---AspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 399
Db 753 GATTCAGTCAAGTATGAGGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
Qy 400 -----GluLysAlaLysLysAenAlaAenLysProLeuLeuAspGluIle 414
Db 813 CCAGCTGGAGATGTGGAGAGAGGCTCTCTGCGGGTTCGCTATGGGGAGGTC 872
Qy 415 ValProValTyrArgArgAspCysHisGluGluValTyrAlaGlySerHisGlnTyrPro 434
Db 873 ATGCCTGTGTACCGCGGACAGCCACCGAGAGCTGCGAGCTGGCAGCATGACTACCT 932
Qy 435 GlyArgGlyValTyrLeuLeuLysPheAspAenSerTyrSerLeuTrpArgSerLysSer 454
Db 933 GGTGAGGCACTCTACCTGCTCAAGTTCGACAACTCTCTACTCTCTGCTGGGCAAGACT 992
Qy 455 ValTyrTyrArgValTyrTyrThr 462
Db 993 CTCTACTTCCACATCTACTACTACAC 1016

```

RESULT 4

[illegible]







```

QY 215 LeuGlnGluGlnHisTyrGln-----GlnTyrLysHisGlnAlaGlu 228
Db 1892 TTAGAGGAGAACCGGCTCGCGATGGAAGAGGAGGAGCAGCAGCTCGCGATGAGGAAGAA 1951
QY 229 Gln-----Gln-----ThrGlnProAla 233
Db 1952 GAACGGAGAGAAAGGAGCTGGAGTCCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2011
QY 234 GlnGlnGlnAlaAlaLeu-----GlnLysGlnGlnGluValValMetAlaGly 249
Db 2012 CAGCAGCAAGAGGCTCTCGGAGGTTGCAGCAGCAGCAGCAGCAACACACAGCTGGCCGAG 2071
QY 250 AlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSerVal 269
Db 2072 ATGAAGCTTCTCTTCTTCAAC-GTGGGGCCAGCAGCTCCAATCAACACAGCATGTCACTG 2130
QY 270 AsnGlyGlnAla-----LysThrHisThrGluAsn 279
Db 2131 CCAGGCCAGCTGTCTGGTGGCTGAAATCCAAAACCTAGAGAAAGAAACGAAACGGCAGCT 2190
QY 280 SerGluLysValLeuGluPro--GluAlaAlaGluGlu-AlaLeuGluAsnGlyProLy 298
Db 2191 TCAGAGAGCAAGGCGCCAGCAGAGGAGTGTGATGAAGCTCTTCAGCAGCAGCAGCA 2250
QY 298 sAspSerLeuProVal-----IleAlaAlaProSerMetTrpThrAr 312
Db 2251 ACAGCAACAGCAGAACTCTCAGGTTGGGGGAATGTCAGCAAACTTCAGTACCAGAA 2310
QY 312 gProGlnIleLysAspPheLysGluLysIleArg-----Gln-AspAlaAsp 328
Db 2311 ATCTCTCTCGAGATCCAGCAGAGGAGGCGCAAAATGCAAAAGCAGCAGCAGCA 2370
QY 328 erValIleThrValArgArgGlyGluValValValArgValProThr----- 344
Db 2371 GCAGCAACACAGCAACCAACAGAGCTCGTAACATACCGCATTCACCTGCACACCAG 2430
QY 345 -----HisGluG 347
Db 2431 CATTTGGGAATCTGTTTGGGCTCTATAATACTGTTCTCTTAACAGTGGCGCATCTGA 2490
QY 347 ludlySerTyrLeuPheTrpGluPheAla---ThrAspSerTyrAspIleGlyPheGly 366
Db 2491 CCTAGTCAGTAGTATT-TGAGTAATGCTGACACTAAAACTCAACATGGGATTC--- 2545
QY 366 alTyrPheGluTTrpThrAspSer----- 373
Db 2546 -----TGGGATGATGCAGTGAAGAGGTGGGACCTAGGAATTCACAAATAAAA 2594
QY 374 -----ProAsnAlaAlaVal-----SerValHisValSerGluSerSerAspGluGluG 390
Db 2595 ATAAAAACAACGCCAGTCTCAGTAATCTGTAGTGTGTCTAACCAGCAGCAATAGAAG 2654
QY 390 ludGluGluGluAsnValThrCysGluGluLysAlaLysLysAsnAlaAsnLys 408
Db 2655 TAGAAGAAGAAAAAGTTGCTG-----AAGCTCTTTTCCAGGAGTAATAAAA 2701

```

## RESULT 10

```

US-10-820-583A-18
; Sequence 18, Application US/10820583A
; Publication No. US20040242461A1
; GENERAL INFORMATION:
; APPLICANT: Schneider, Michael
; APPLICANT: Oh, Hidemasa
; TITLE OF INVENTION: Modulators of Telomere Stability
; FILE REFERENCE: HO-P02673US1
; CURRENT APPLICATION NUMBER: US/10/820, 583A
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: US 60/461,095
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18

```

```

; LENGTH: 3792
; TYPE: DNA
; ORGANISM: HUMAN
US-10-820-583A-18

```

## Alignment Scores:

```

Pred. No.: 9,28e-13 Length: 3792
Score: 229.00 Matches: 76
Percent Similarity: 48.45% Conservative: 49
Best Local Similarity: 29.46% Mismatches: 78
Query Match: 9.57% Indels: 56
DB: 20 Gaps: 10

```

```
US-09-762-594-7 (1-463) x US-10-820-583A-18 (1-3792)
```

```

QY 115 ArgIleGluLysGluGluGluGluLysArgArgLysAlaGluGluGluArgGlnArg 134
Db 1192 CGGATTGAGCAGCAGAAAGAACAGAGGCGCTAGAGAGCAACAAAGAGAGAGCGG 1251
QY 135 Glu-----GluGluGluArgGluArgLeuGlnLysGluGluLysArgLys 150
Db 1252 GAAGCTAGAGGCGAGCAGGAACTGAAACAGCGAAGGAGAGAACAGAGAAAGAGCGGT 1311
QY 151 ArgGluGluGluAspArgLeuArgGluGluGluGluArgArgIleGluGluGlu 170
Db 1312 CTAGAGGAGTTGAGAGAGAGCGCAAGAAAGAGAGAGAGAGAGAGCGGCGAGAAGAA 1371
QY 171 ArgLeuArgLeuGlnGlnLysGlnGlnIleMetAlaAlaLeuAsn----- 186
Db 1372 AAGAGGAGAGTTGAAAGAGAACAGAGATATATCAGCGCAGCAGCTAGAGAGGAGCAGCG 1431
QY 187 -----SerGlnThrAlaValGlnPheGln----- 194
Db 1432 CACTTGGAGTCTCTCAGCAGCAGCTGCTCCAGGAGAGGCCATGTTACTGGAGTGCAG 1491
QY 195 -----GlnTyrAlaGlnGlnTyrProGlyAsnTyrGluGlnGln 208
Db 1492 TGGCGGAGATGAGGAGCAGCGGAGGAGAGAGGCTCCAGAGGCGAGTTGCAACAGAA 1551
QY 209 GlnIleLeuIleArgGlnGlnGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGlu 228
Db 1552 CAAGCATATCTCTCTCTACAGCATGACCATAGGAGGCGCAGCCGCGCAGCAGCTCGCAG 1611
QY 229 GlnThrGlnProAlaGlnGln-AlaAlaLeuGlnLysGlnGlnValValMetAl 248
Db 1612 CAGCCGCCACCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1648
QY 248 aGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268
Db 1649 ----GCTTCCATGCTCCCGAGCCCAAGCCACTACGAGCCT-GCTGACCGGCGGAGA 1703
QY 268 rValAsnGlyGln---AlaLysThrHisThrGluAsnSerGlu----- 281
Db 1704 GSTGGAAGATAGATTTTAGGAAACCTAACACAGCTCCCTCTGAAGCCGAGTCTTAAGCAGAC 1763
QY 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLy 298
Db 1764 AGCAGAGTATTGGAGCCACAGTGCCTTCCGATCAGAGTCTTTTCCATATGGCAATC 1823
QY 298 sAspSerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
Db 1824 CGAGTCTGTGATCCCGCCTCGAGAGCAGCG-----GAGCCACAGGTT----- 1869
QY 317 pPheLysGluLysIleArgGlnAspAlaAspSerValIleThrValArgArg 334
Db 1870 -----CCTGTGAGAAACACATCTCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 1911

```

## RESULT 11

```

US-09-291-417-9
; Sequence 9, Application US/09291417A
; Publication No. US20030050230A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY

```

```

; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 240/300
; CURRENT APPLICATION NUMBER: US/09/291,417A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: US 60/081,784
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 3798
; TYPE: DNA
; ORGANISM: Mammalian (Human) ZC1
; US-09-291-417-9

Alignment Scores:
Pred. No.:          9.3e-13          Length:          3798
Score:              229.00           Matches:         76
Percent Similarity: 48.45%           Conservative:    49
Best Local Similarity: 29.46%         Mismatches:     78
Query Match:        9.57%            Indels:         56
DB:                  10              Gaps:           10

US-09-762-594-7 (1-463) x US-09-291-417-9 (1-3798)

QY 115 ArgfileGluLysGluGluGluLysArgArgLysAlaGluGluLysArgGlnArg 134
Db 1198 CGGATTGAGCAGCAAGAAAGACAGAGCGCGCTAGAGAGCAACAAAGAGAGAGCGG 1257

QY 135 Glu-----GluGluGluArgGluArgLysGluGluGluLysGluLysArgLys 150
Db 1258 GAAGCTAGAGGCGAGCAACGCTGAACAGCAGAGAGAGAGAGAGAGAGAGCGCT 1317

QY 151 ArgGluGluLysArgLeuArgArgGluGluGluGluGluGluGluGluGluGlu 170
Db 1318 CTAGAGAGAGTTGGAGAGAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1377

QY 171 ArgLeuArgLeuGluGlnGlnLysGlnGlnMetAlaAlaLeuAsn----- 186
Db 1378 AAGAGGAGAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1437

QY 187 -----SerGlnThrAlaValGlnPheGln----- 194
Db 1438 CACTTGGAAAGTCCTTCAGCAGCAGCTGCTCCAGGAGCAGCGCATGTTACTGGAGTCCGA 1497

QY 195 -----GlnTyralaAlaGlnGlnTyrrProGlyAsnTyrrGluGlnGln 208
Db 1498 TGGCGGGAGATGGAGAGCACCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557

QY 209 GlnIleLeuIleArgGlnGluGlnGlnHisTyrrGlnGlnTyrrLysHisGlnAlaGlu 228
Db 1558 CAAGCATATCTCTGCTCTACAGCATGACCATAGAGAGAGAGAGAGAGAGAGAGAG 1617

QY 229 GlnThrGlnProAlaGlnGln-AlaAlaLeuGlnLysGlnGlnValValMetAl 248
Db 1618 CAGCCGCCACCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1654

QY 248 aGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268
Db 1655 ----GTTTCATGCTCCCGGCCCAAGCCCACTACAGAGCCT-GCTGACCGAGCGCGAGA 1709

QY 268 rValAsnGlyGln---AlaLysThrHisThrGluAsnSerGlu----- 281
Db 1710 GGTGGAAGATAGATTATTAGGAAACTAACCACAGCTCCCTGGAAGCCAGCTCTTAAGCAGAC 1769

QY 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProly 298
Db 1770 AGCAGAGATTGGAGCCACCAAGTGCCTTCCCGATCAGAGTCTTTTCCATGGCAATC 1829

QY 298 sAspSerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
Db 1830 CGAGTCTGTGCATCCCGCTGCAGAGACAGCG-----GAGCCACAGGTT----- 1875

```

```

QY 317 pPhelysGluLysIleArgGlnAspAlaAspSerValIleThrValArgArg 334
Db 1876 -----CCTGTGAGAAACAACATCTCGCTCCCTGTTCTGTCCCGTCGA 1917

RESULT 12
US-10-725-329-9
; Sequence 9, Application US/10725329
; Publication NO. US20040224323A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/10/725,329
; CURRENT FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US/09/688,188B
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3798
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-725-329-9

```

```

Alignment Scores:
Pred. No.:          9.3e-13          Length:          3798
Score:              229.00           Matches:         76
Percent Similarity: 48.45%           Conservative:    49
Best Local Similarity: 29.46%         Mismatches:     78
Query Match:        9.57%            Indels:         56
DB:                  10              Gaps:           10

```

US-09-762-594-7 (1-463) x US-10-725-329-9 (1-3798)

```

QY 115 ArgfileGluLysGluGluGluLysArgArgLysAlaGluGluGluArgGlnArg 134
Db 1198 CGGATTGAGCAGCAAGAAAGACAGAGCGCGCTAGAGAGCAACAAAGAGAGAGCGG 1257

QY 135 Glu-----GluGluGluArgGluArgLeuGluGlnLysGluGluLysArgLys 150
Db 1258 GAAGCTAGAGGCGAGCAACGCTGAACAGCAGAGAGAGAGAGAGAGAGAGAGCGCT 1317

QY 151 ArgGluGluLysArgLeuArgArgGluGluGluGluGluGluGluGluGluGlu 170
Db 1318 CTAGAGAGAGTTGGAGAGAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1377

QY 171 ArgLeuArgLeuGluGlnGlnLysGlnGlnMetAlaAlaLeuAsn----- 186
Db 1378 AAGAGGAGAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1437

QY 187 -----SerGlnThrAlaValGlnPheGln----- 194
Db 1438 CACTTGGAAAGTCCTTCAGCAGCAGCTGCTCCAGGAGCAGCGCATGTTACTGGAGTCCGA 1497

QY 195 -----GlnTyralaAlaGlnGlnTyrrProGlyAsnTyrrGluGlnGln 208
Db 1498 TGGCGGGAGATGGAGAGCACCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557

QY 209 GlnIleLeuIleArgGlnGluGlnGlnHisTyrrGlnGlnTyrrLysHisGlnAlaGlu 228
Db 1558 CAAGCATATCTCTGCTCTACAGCATGACCATAGAGAGAGAGAGAGAGAGAGAGAG 1617

QY 229 GlnThrGlnProAlaGlnGln-AlaAlaLeuGlnLysGlnGlnValValMetAl 248
Db 1618 CAGCCGCCACCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1654

```

```
QY 248 aGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268
Db 1655 ----GCTTCATGCTCCGAGCCCAAGCCACTACGAGCCT-GCTGACCGAGCGGAGA 1709
QY 268 rValAsnGlyGln---AlaLysThrHisThrGluAsnSerGlu----- 281
Db 1710 GGTGGAGATAGATTAGGAAACTAACACACACTCCCTGAAGCCAGCTCTTAAGCAGAC 1769
QY 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLy 298
Db 1770 AGGCAGAGTATTGGAGCCACAGTGCTTCCCGATCAGAGTCTTTTCCAAATGGCAACTC 1829
QY 298 sAspSerLeu---ProValleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
Db 1830 CGAGTCTGTGCATCCCGCCCTGCAGAGACCAGCG-----GAGCCACAGGTT----- 1875
QY 317 pPhelYsGluLysIleArgGlnAspAlaAspSerValIleThrValArgArg 334
Db 1876 -----CCTGTGAGAACAAACATCTCGCTCCCTGTTCTGTCTCCGTCGA 1917

RESULT 13
US-10-725-121-9
; Sequence 9, Application US/10725121
; Publication No. US20050142625A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/10/725,121
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US/09/688,188B
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3798
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-725-121-9

Alignment Scores:
Pred. No.: 9.3e-13 Length: 3798
Score: 229.00 Matches: 76
Percent Similarity: 48.45% Conservative: 49
Best Local Similarity: 29.46% Mismatches: 78
Query Match: 9.57% Indels: 56
DB: 22 Gaps: 10

US-09-762-594-7 (1-463) x US-10-725-121-9 (1-3798)

QY 115 ArgIleGluLysGluGluGluLysArgArgLysAlaGluGluGluArgGlnArg 134
Db 1198 CGAATTGACAGCAAGAAAGACAGAGCGCGCTAGAGAGCAACAAAGAGAGAGCGG 1257
QY 135 Glu-----GluGluGluArgGluArgLysGluLysGluLysArgLys 150
Db 1258 GAAGCTAGAGGCGAGCAGCAAGCTGAACAGCGAGGAGAGACAAAGAAAGAGCGCT 1317
QY 151 ArgGluGluAspArgLeuArgGluGluGluGluArgArgArgIleGluGluGlu 170
Db 1318 CTAGAGGAGTTGGAGAGAGCGCAAGAAAGAGAGAGAGAGAGAGAGCGGAGAGAA 1377
QY 171 ArgLeuArgLeuGlnGlnLysGlnGlnIleMetAlaLeuAsn----- 186
Db 1378 AAGAGGAGAGTTGAAAGAGAAACAGAGATATATACGCGCAGCTAGAGAGAGAGCGG 1437
QY 187 -----SerGlnThrAlaValGlnPheGln----- 194
```

```
Db 1438 CACTTGGAAAGTCTCTCAGCAGCAGCTGCTCCAGGAGCAGCCATGTTACTGGAGTCCCGA 1497
QY 195 -----GlnTyrAlaAlaGlnGlnTyr-ProGlyAsnTyrGluGlnGln 208
Db 1498 TGGCGGGAGATGAGGAGCACCAGGAGGAGAGAGCTCCAGAGGAGGAGTTCGAACAGAA 1557
QY 209 GlnIleLeuIleArgGlnLeuGlnGlnHisTyrGlnGlnTyrLysHisGlnAlaGlu 228
Db 1558 CAAGCATATCTCTGCTCTACAGCATGACCATAGGAGCCGACCCGAGCAGCACTCGCAG 1617
QY 229 GlnThrGlnProAlaGlnGlnGln-AlaAlaLeuGlnLysGlnGlnGlnValMetAl 248
Db 1618 CAGCCGCCACCAACCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1654
QY 248 aGlyAlaSerLeuProAlaSerSerLysValAsnThrAlaGlyAlaSerAspThrLeuSe 268
Db 1655 ----GCTTCATGCTCCGAGCCCAAGCCACTACGAGCCT-GCTGACCGAGCGGAGA 1709
QY 268 rValAsnGlyGln---AlaLysThrHisThrGluAsnSerGlu----- 281
Db 1710 GGTGGAGATAGATTAGGAAACTAACACACACTCCCTGAAGCCAGCTCTTAAGCAGAC 1769
QY 282 ----LysValLeuGluPro-----GluAlaAlaGluGluAlaLeuGluAsnGlyProLy 298
Db 1770 AGGCAGAGTATTGGAGCCACAGTGCTTCCCGATCAGAGTCTTTTCCAAATGGCAACTC 1829
QY 298 sAspSerLeu---ProValleAlaAlaProSerMetTrpThrArgProGlnIleLysAs 317
Db 1830 CGAGTCTGTGCATCCCGCCCTGCAGAGACCAGCG-----GAGCCACAGGTT----- 1875
QY 317 pPhelYsGluLysIleArgGlnAspAlaAspSerValIleThrValArgArg 334
Db 1876 -----CCTGTGAGAACAAACATCTCGCTCCCTGTTCTGTCTCCGTCGA 1917

RESULT 14
US-10-494-940-21
; Sequence 21, Application US/10494940
; Publication No. US20050069886A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: REGULATED PROSTATE CANCER GENES
; FILE REFERENCE: OGT 9U 803 PCT
; CURRENT APPLICATION NUMBER: US/10/494,940
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/331,042
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/331,041
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/340,251
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/344,791
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 3612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (312)..(3611)
; OTHER INFORMATION:
US-10-494-940-21

Alignment Scores:
Pred. No.: 9.77e-13 Length: 3612
Score: 228.50 Matches: 95
Percent Similarity: 46.81% Conservative: 59
Best Local Similarity: 28.88% Mismatches: 95
Query Match: 9.54% Indels: 81
DB: 21 Gaps: 13
```



Qy 300 SerLeu---ProValIleAlaAlaProSerMetTrpThrArgProGlnIleLysAspPhe 318  
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::  
Db 1874 TCTGTGCATCCCGCCCTGCAGACACGCG-----GAGCCACAGGTT----- 1915  
  
Qy 319 LysGluLysIleArgGlnAspAlaAspSerValIleThrValArgArg 334  
|||::: |||::: |||::: |||::: |||::: |||::: |||:::  
Db 1916 -----CCTGTGAGAACACATCTCGCTCCCTGTTCTGTCCCGTCGA 1957

Search completed: August 23, 2005, 02:22:57  
Job time : 982 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 09:05:19 ; Search time 1019 Seconds  
(without alignments)  
9302.262 Million cell updates/sec

Title: US-09-762-594-2  
Perfect score: 1459  
Sequence: 1 gaattcgccgcgcgcgcac.....catgacgttaatttccttt 1459

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues  
Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA: \*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq: \*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq: \*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq: \*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq: \*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq: \*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq: \*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq: \*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq: \*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq: \*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1073.2	73.6	1481	21	US-10-998-342-9
2	268.2	18.4	278	19	Sequence 9, Appli
3	204.4	14.0	934	9	Sequence 327, App
4	150.8	10.3	3547	17	Sequence 33, Appl
5	150	10.3	882	18	Sequence 832, App
6	113.4	7.8	527	9	Sequence 238, App
7	110.6	7.6	554	9	Sequence 32, Appl
C					Sequence 13381, A

C	8	110.6	7.6	210204	20	US-10-723-860-1746	Sequence 1746, Ap
C	9	109.2	7.5	306	9	US-09-864-761-30285	Sequence 30285, A
C	10	91.8	6.3	535	20	US-10-357-930-55215	Sequence 55215, A
C	11	88.8	6.1	812	17	US-10-012-697-1343	Sequence 1343, Ap
C	12	78.4	5.4	486	17	US-10-465-211-13	Sequence 13, Appl
C	13	77.2	5.3	684	17	US-10-012-697-1045	Sequence 1045, Ap
C	14	74.2	5.1	628	16	US-10-029-385-22859	Sequence 22859, A
C	15	74.2	5.1	3489	11	US-09-894-273-1	Sequence 1, Appli
C	16	74.2	5.1	3489	15	US-10-294-804-1	Sequence 1, Appli
C	17	74.2	5.1	3489	20	US-10-194-046-1	Sequence 1, Appli
C	18	72.6	5.0	1926	15	US-10-294-804-3	Sequence 3, Appli
C	19	72.6	5.0	1926	20	US-10-194-046-3	Sequence 3, Appli
C	20	72.6	5.0	8705	15	US-10-291-230-14	Sequence 14, Appl
C	21	72.6	5.0	8705	15	US-10-291-249-14	Sequence 14, Appl
C	22	72.6	5.0	8705	18	US-10-273-678-16	Sequence 16, Appl
C	23	72.6	5.0	9482	21	US-10-888-961-4	Sequence 4, Appli
C	24	72.6	5.0	9600	16	US-10-278-751-1	Sequence 1, Appli
C	25	72.6	5.0	10233	16	US-10-050-898-283	Sequence 283, App
C	26	72.6	5.0	10285	16	US-10-050-902-283	Sequence 283, App
C	27	72.6	5.0	10330	19	US-10-656-269-24	Sequence 24, Appl
C	28	72.6	5.0	10477	19	US-10-656-269-22	Sequence 22, Appl
C	29	72.6	5.0	10516	19	US-10-656-269-20	Sequence 20, Appl
C	30	72.6	5.0	10561	19	US-10-656-269-18	Sequence 18, Appl
C	31	72.6	5.0	10615	19	US-10-656-269-45	Sequence 45, Appl
C	32	72.6	5.0	10774	19	US-10-656-269-23	Sequence 23, Appl
C	33	72.6	5.0	10921	19	US-10-656-269-21	Sequence 21, Appl
C	34	72.6	5.0	10961	19	US-10-656-269-19	Sequence 19, Appl
C	35	72.6	5.0	11006	19	US-10-656-269-17	Sequence 17, Appl
C	36	72.6	5.0	11059	19	US-10-656-269-46	Sequence 46, Appl
C	37	72.6	5.0	11924	19	US-10-678-816-7	Sequence 7, Appli
C	38	72.6	5.0	12242	19	US-10-678-816-6	Sequence 6, Appli
C	39	72.6	5.0	16080	24	US-11-089-918-48	Sequence 48, Appl
C	40	72.6	5.0	16080	24	US-11-090-865-48	Sequence 48, Appl
C	41	72.6	5.0	16080	24	US-11-069-377-48	Sequence 48, Appl
C	42	72.6	5.0	16080	24	US-11-090-872-48	Sequence 48, Appl
C	43	72.2	4.9	102374	13	US-10-087-192-667	Sequence 667, App
C	44	71.4	4.9	68233	17	US-10-034-650-31	Sequence 31, Appl
C	45	70.4	4.8	2418	19	US-10-437-963-73494	Sequence 73494, A

ALIGNMENTS

RESULT 1  
US-10-998-342-9  
; Sequence 9, Application US/10998342  
; Publication No. US20050106672A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: LAL, Preeti  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: GORGONE, Gina A.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: PATTERSON, Chandra  
; TITLE OF INVENTION: NEUROTRANSMISSION ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0551 PCT  
; CURRENT APPLICATION NUMBER: US/10/998,342  
; CURRENT FILING DATE: 2004-11-29  
; PRIOR APPLICATION NUMBER: US/09/720,530  
; PRIOR FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 60/091,667  
; PRIOR FILING DATE: 1998-07-02  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 1481  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte Clone No: 99868  
US-10-998-342-9

Query Match 73.6%; Score 1073.2; DB 21; Length 1481;  
Best Local Similarity 86.9%; Pred. No. 9.6e-307;  
Matches 1180; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 97 AGTAAAGATGGCAAGCCCTTTCATCCACTTATGAAGAAAACCTGAAGTTCGTGCACCT 156  
DB 70 AGAAAAGATGGCAAGCAATTCATCCAACTTATGAAGAAAATTAAGCTTGTGGCACT 129  
QY 157 GCATAAGCAAGTTCTTTTGGGCCCATATAACCCAGACACGTCCTCCCTGAGGTTCGATTCCT 216  
DB 130 GCATAAGCAAGTTCTTATGGGCCCATATATATCCAGACACATTGTCTCCAGGTTGGATTCCT 189  
QY 217 TGATGTGTTGGGAATGATAGGAGGAGAGAAATGGGCAGCTCTGGGAAACATGTCCAAAGGA 276  
DB 190 TGATGTGTTGGGGAATGACAGGAGGAGAGAAATGGGCAGCCCTGGGAAAACATGTCTAAAGA 249  
QY 277 GATGCCATGGTAGAGTTGTGAAGCTTCTAAATAAGTGTGTCTCTCTCTCTCTCTCTCTCTCT 336  
DB 250 GATGCCATGGTAGAGTTGTCAAGCTCTTAAATAGGTGTGTCTCTCTCTCTCTCTCTCTCTCT 309  
QY 337 TGTTCGTGCCACAGAAATAGAGAGGAGAGAAAGAGAGAGAAAGAGAAAGCGGAGGAGGA 396  
DB 310 TGTTCGTGCCACAAATAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 369  
QY 397 GCGAAGCGACGCTGAAGAGGAGAGAAACGAGAGCGGCTGCAAAAGGAGAGAGAGAGCGGAA 456  
DB 370 GCGAAGCGGCGTGAAGAGGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 429  
QY 457 GCGAGAGGAGAGACCGGCTGAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516  
DB 430 GAGAGAGAGAGAGAGAGAGGCTTCACGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 489  
QY 517 GAGGCTTCGCTCGAAACAGCAAGCAGCAGATTAATGGCAGCTTTAAACTCGCAGACTGC 576  
DB 490 AAGGCTTCGCTCGAGCAGCAAGCAGCAGATTAATGGCAGCTTTAAACTCCGAGACTGC 549  
QY 577 CGTGCAATTCACGAGTATGACGCCAGCAGTATCCAGGGAATACGAAACACAGCAGAT 636  
DB 550 CGTGCAATTCACGAGTATGACGCCAACAGTATCCAGGGAATACGAAACAGCAGCAAT 609  
QY 637 TCTCATCCGACCTGCAGGAGCAGCACTATCAGCAGTATTAACACACGAGCAGGCAAC 696  
DB 610 TCTCATCCGACCTGCAGGAGCAACACTATCAGCAGTATTAACACACGAGCAGTGTATCAAGT 669  
QY 697 CCAACTGCAACAACAGCAGCAGCATTACAGAAACAGCAGAGTATGATGCTGGGCG 756  
DB 670 CCAAGTTCACAGCAACAGCAGCAGTATTAACAGAAACAGCAGAGTATGATGCTGGGTC 729  
QY 757 ATCAATTCGCTGCAATCAAAAGGTGAACACAGCTGGAGCAAGTGATACACTGTCACTTAA 816  
DB 730 TTTCTTTGCTTACATCATCAAAAGTGAATGCAACTGTACCAAGTAAATATGATGTCACTTAA 789  
QY 817 TGACAGGCGCAAAACCAACACTGAAATTCGGAAGAGTCTTTCAGCAGGAGAGCTGCAGA 876  
DB 790 TGACAGGCGCAAAACCAACACTGCACTGCAAGTCTCGAAAGAGAACTGGAACCAAGAGCTGCAGA 849  
QY 877 AGAAGCCTTCGAAATAGGACAAAGACTCTCTTCAGTGTATTCAGCTCCATCCATGTG 936  
DB 850 AGAAGCCTTCGAGATAGGACAAAGAACTCTCTTCAGTGTATTCAGTATAGCAGCTCCATGTG 909  
QY 937 GACAAGACCAAAATCAAGACTTTAAAGAGAAAGATTCGCGAGGATGCAGATTCGTGAT 996  
DB 910 GACAAGACCTCAGATCAAGACTTTCAAGAGAGAAATTCAGCAGGATGCAGATTCGCTGAT 969  
QY 997 TACAGTACGTGAGGAGAGTCTGACCGTCCAGTCCGAGTCCGAGCTCATGAGAGAGATCATTA 1056  
DB 970 TACAGTGGGCGGAGGAGAGTGTGTCTGTCTGAGTACCCACCATGAAAGAGGATCATTA 1029  
QY 1057 CCTATTTCGGAATTTGCCACAGACAGTATGACATTCGAGTTGGGTTTATTTTCAATG 1116

DB 1030 TCTCTTTTGGGAATTTGCCACAGACAAATTAACACATTTGGGTTTGGGCTGTATTTTGAATG 1089  
QY 1117 GACAGACTCTCCAAATGCTGCTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATG 1176  
DB 1090 GACAGACTCTCCAAACACTGCTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATG 1149  
QY 1177 GGAGGAGGAGAAATATGTCACCTTGTGAAGAAAAGCAAAAGAGAGAGAGAGAGAGAGAGAG 1236  
DB 1150 GGAGGAGGAGAAATATGTCGTTGTGAAGAAAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAG 1209  
QY 1237 GCTGATGAGATTTGATCTGTGTACCGGCGGAGCTGTACGAGGAGAGTATATGAGGAGAG 1296  
DB 1210 GCTGATGAGATTTGCTGTGTACCGGCGGAGCTGTATGAGGAGAGTGTATGCTGCTGCTG 1269  
QY 1297 CCACAGATATCCAGGAGGAGAGTCTATCTCTCAAGTTTGATTAATTCCTACTCTCTG 1356  
DB 1270 CCATCAATATCCAGGAGGAGAGTCTATCTCTCAAGTTTGATTAATTCCTACTCTCTG 1329  
QY 1357 GAGGTCCAGTCCGCTACTACAGAGTCTATTAATTAATTAATTAATTAATTAATTAATTAAT 1416  
DB 1330 GCGGTCAAAATCAGTCTACTACAGAGTCTATTAATTAATTAATTAATTAATTAATTAAT 1389  
QY 1417 CCGGAGTCCAGGTTTGAGCACAACATGAGCTTTAAATTT 1454  
DB 1390 CTGGAGTCTAGGTTGGGCGAGAGATGACATTTAATTT 1427

RESULT 2  
US-10-469-285-327/c  
; Sequence 327, Application US/10469285  
; Publication No. US20040126776A1  
; GENERAL INFORMATION:  
; APPLICANT: LION Bioscience AG  
; TITLE OF INVENTION: Gene library  
; FILE REFERENCE: L 1541  
; CURRENT APPLICATION NUMBER: US/10/469,285  
; CURRENT FILING DATE: 2003-08-28  
; NUMBER OF SEQ ID NOS: 840  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 327  
; LENGTH: 278  
; TYPE: DNA  
; ORGANISM: Mus Musculus  
US-10-469-285-327

Query Match 18.4%; Score 268.2; DB 19; Length 278;  
Best Local Similarity 98.9%; Pred. No. 1.4e-68;  
Matches 270; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1187 GAAATGTCACCTGTGACGCGGAGCTGTACGAGGAGTATATGACGAGCAGCACCACAGTAT 1246  
DB 277 GAAATGTCACCTGTGAGAGAAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218  
QY 1247 ATTGTACCTGTGTACGCGGAGCTGTACGAGGAGTATATGACGAGCAGCACCACAGTAT 1306  
DB 217 ATTGTACCTGTGTACGCGGAGCTGTACGAGGAGTATATGACGAGCAGCACCACAGTAT 158  
QY 1307 CCAGGAGGAGTCTATCTCTCAAGTTTGATTAATTCCTACTCTCTGTGAGGTCGAAG 1366  
DB 157 CCAGGAGGAGTCTATCTCTCAAGTTTGATTAATTCCTACTCTCTGTGAGGTCGAAG 98  
QY 1367 TCCGCTACTACAGTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1426  
DB 97 TCCGCTACTACAGTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 38  
QY 1427 GGGTTGAGCACAACATGAGCTTTAATTTTCTTT 1459  
DB 37 GGGTTGAGCACAACATGAGCTTTAATTTTCTTT 5

RESULT 3  
US-09-866-562-33  
; Sequence 33, Application US/09866562













